

Journal of Official Statistics, Vol. 36, No. 2, 2020, pp. 361–378, http://dx.doi.org/10.2478/JOS-2020-0019

# Measuring the Sustainable Development Goal Indicators: An Unprecedented Statistical Challenge

Steve MacFeely<sup>1</sup>

In March 2017, the United Nations (UN) Statistical Commission adopted a measurement framework for the UN Agenda 2030 for Sustainable Development, comprising of 232 indicators designed to measure the 17 Sustainable Development Goals (SDGs) and their respective 169 targets. The scope of this measurement framework is so ambitious it led Mogens Lykketoft, President of the seventieth session of the UN General Assembly, to describe it as an 'unprecedented statistical challenge'.

Naturally, with a programme of this magnitude, there will be foreseen and unforeseen challenges and consequences. This article outlines some of the key differences between the Millennium Development Goals and the SDGs, before detailing some of the measurement challenges involved in compiling the SDG indicators, and examines some of the unanticipated consequences arising from the mechanisms put in place to measure progress from a broad political economy perspective.

*Key words:* 2030 Agenda; unintended consequences; national statistical systems; administrative data.

# 1. Introduction

The 2030 Agenda for Sustainable Development and the Sustainable Development Goals (SDG) Global Indicator Framework (GIF) was adopted by the United Nations (UN) Statistical Commission in March 2017 (UN Statistical Commission 2017) and subsequently adopted by the UN General Assembly in July 2017 (UN General Assembly 2017). The framework comprises 232 statistical indicators designed to measure the seventeen 2030 Agenda goals and their respective 169 targets. The aim of the GIF is to provide good quality, verifiable evidence on progress towards achieving the 2030 Agenda. However, populating those indicators and providing that evidence poses enormous challenges. So much so, it led Mogens Lykketoft, President of the seventieth session of the UN General Assembly, to describe it as an 'unprecedented statistical challenge' (Lebada 2016).

This article outlines some of the measurement challenges involved in compiling the SDG indicators, some of the tensions that have arisen during the process to date, and also examines some of the unanticipated consequences arising from the mechanisms put in place to measure progress. By explaining, very briefly, some of the key differences between the Millennium Development Goals (MDGs) and the SDGs, and the political circumstances

<sup>1</sup> United Nations Conference on Trade and Development, Palais des Nations, CH-1211 Geneva 10, Switzerland. Email: Steve.macfeely@un.org in which the GIF came to life, the article attempts to explain why this project is indeed an unprecedented statistical challenge.

The remainder of this article is presented in nine sections. The next section identifies some of the most important differences between the SDGs and their predecessor, the MDGs. The following two sections outline some of the challenges and tensions that have emerged in measuring the SDGs. Sections 5 and 6 explain how the Inter-agency and Expert Group on SDG Indicators (IAEG-SDG) has classified the SDG indicators into tiers, and details also some of the cost estimates for compiling SDG indicators put forward by several commentators. Section 7 outlines some of the unintended consequences that are emerging from the process. The penultimate section presents a discussion of some issues arising, before the article is concluded with some recommendations.

#### 2. From MDGs to SDGs

At the beginning of 2016, the UN SDGs replaced the MDGs, which had been in place since the turn of the century. Although both sets of goals describe an aspirational road map for global development, the SDGs came about through a profoundly different process than the MDGs, which was essentially a distillation of the major agreements from the main development conferences of the 1990s, (such as the World Summit for Children 1990; the UN Conference on Environment and Development 1992; the World Conference on Human Rights 1993; the International Conference on Population and Development 1994; the World Summit for Social Development 1995; the Fourth World Conference on Women 1995; the Second UN Conference on Human Settlements Habitat II 1996; the World Food Summit 1996). These agreements were compiled by the UN Secretariat and reflected in the UN Secretary General's Millennium Report, We the Peoples: The Role of the United Nations in the 21st Century (Annan 2000), which outlined the challenges for development in a globalised world. At the fifty-fifth General Assembly, designated the 'Millennium Summit', 189 Member States adopted the Millennium Declaration (UN 2000). This Declaration committed nations to reduce extreme poverty by 2015. The following year, in August 2001, the UN Secretariat published the final set of eight MDGs. It was described by Hulme (2009, 4) as 'the world's biggest promise.'

The SDGs, by contrast, emerged from the shadow of the MDGs, which had been criticised for pushing a donor-driven agenda, excluding any discourse critical of the Washington Consensus and not fully reflecting the will or views of peoples or governments. From the outset, the SDG process aimed to create a people-centred development agenda. To do so, an unprecedented global consultation was undertaken. Following three years of consultation and negotiation, involving thousands of people, *Transforming Our World: The 2030 Agenda for Sustainable Development* (UN 2015a) was formally adopted by 193 heads of government, including 150 heads of state on 25 September 2015. The 2030 Agenda adopted a broad view of development, one that encompassed not just ending extreme poverty and eradicating hunger, but one that aspires to foster global prosperity in an economically and environmentally sustainable and equitable way. The 17 SDGs and their 169 targets would be 'action oriented, global in nature and universally applicable' (UN, 2013, 4), and were described by Ban Ki-moon (UN 2015b), former Secretary General of the UN, as the 'to do list for planet and people'.

#### 3. Some Challenges in Measuring the SDGs

From a statistical perspective, the implications of the 2030 Agenda and the accompanying GIF are enormous. Not only have the number of goals and targets increased considerably compared with the MDGs (The MDGs had 8 goals, 21 targets and 60 indicators, whereas the SDGs have 17 goals, 169 targets and 232 indicators), but so also has the complexity of these targets. The scope of the 2030 Agenda is also far wider than that of its predecessor, attempting to span the full spectrum of development issues, including not only aspects of society, economy and the environment, but also institutional coordination.

A simple illustration of the complexity is available from the report *Data Disaggregation and SDG Indicators: Policy Priorities and Current and Future Disaggregation Plans* (IAEG-SDG, 2019). This matrix details the minimum set of disaggregation required for each indicator. The level of disaggregation varies considerably by indicator. For example, the "minimum required disaggregation dimension" demanded by Target 1.3 (Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable) is: sex; age; employment status; disability status; pregnancy; work-injury victims; and income. For Target 10.2 (By 2030, empower and promote the social, economic and political inclusion of all, irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status) the "minimum required disaggregation dimension" is: sex; age; disability status; race; ethnicity; origin; religion; and other economic or social status.

The first challenge facing statisticians was to clarify what it was they were being asked to measure. This was easier said than done. Deciphering or interpreting exactly what is meant by the agreed text of *Transforming our World: The 2030 Agenda for Sustainable Development* (UN 2015a) was not always straightforward. Lack of clear definitions and inconsistent use of terminology are just some examples of where statisticians, in selecting appropriate indicators, were forced to decide what the targets actually meant. For example, what is meant by 'sustainable'?

Does it just mean environmentally sustainable, or does it also mean economically sustainable, or socially sustainable? Environmentalists will naturally assume it means environmentally sustainable, but economists will equally assume it means economic sustainability. The next question is how long a trend should be exhibited before it can be considered sustainable – will this be the same for economic or environmental variables? What about 'economic stability'? Target 17.13 calls for global macro-economic stability. Although there is no consensus on what this means, it has been agreed it will be measured by a dashboard of indicators. The composition of this dashboard will effectively determine whether the 2030 Agenda adopts an orthodox or heterodox view of the global economy.

What are the 'basic services' or the 'new technologies' referred to in Target 1.4 (By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance) and are they the same in all parts of the world? This might seem like pedantry, but it matters when you are trying to design an appropriate measurement. A plethora of seemingly commonly understood words, such as: access; adverse; adequate; appropriate; basic; benefit; efficient; effective;

informal; infrastructure; integration; promote; resilience; resource; sustainable; and vulnerable caused comprehension problems and challenges of consistent interpretation across the 169 targets, requiring the construction of a SDG ontology (UN Environmental Programme 2015) to make progress.

Another challenge is the lack of priority within complex and sometimes rather muddled targets. This has proven particularly thorny, as statisticians were instructed by their political masters to limit the number of indicators to one indicator per target. Numerate readers will have noted that this guideline was not respected, as 169 targets resulted in 232 indicators. In truth, to measure the targets properly, closer to 500 indicators would probably be required. Take Target 17.19 (By 2030, build on existing initiatives to develop measurements of progress on sustainable development that complement GDP and support statistical capacity building in developing countries) for example. This target combines two completely different and unrelated issues: firstly, the measurement of progress beyond GDP and secondly, supporting statistical capacity-building. This bundling, not uncommon to many targets, poses a dilemma. Which element of the target should be measured? Both are very important, but both are also very complex. The challenge of how to properly measure progress is a highly contentious issue, hotly debated by economists, social scientists, environmentalists and statisticians (MacFeely 2016), and would probably need a whole dashboard of indicators to do justice to this one issue. Equally, the best way to approach statistical capacity-building is also being actively discussed and reassessed (Jütting 2016). The idea that such a cocktail of issues could sensibly be amalgamated into a single indicator is absurd. The Economist (2015b), citing Target 4.7 (By 2030, ensure all learners acquire the knowledge and skills needed to promote sustainable development, including among others through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and non-violence, global citizenship and appreciation of cultural diversity and of culture's contribution to sustainable development) as an example, put it bluntly, simply saying, 'try measuring that.'

Unsurprisingly, indicator 4.7.1 (Extent to which (i) global citizenship education and (ii) education for sustainable development, including gender equality and human rights, are mainstreamed at all levels in: (a) national education policies; (b) curricula; (c) teacher education; and (d) student assessment) has been classified as Tier 3 by the IAEG-SDGs (see Section 5).

Although the scope of the 2030 Agenda is universal and applies to all countries, clearly not all targets are relevant to every country. Striking a balance between national and global demands has proven challenging. For example, Target 3.3 (By 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, waterborne diseases and other communicable diseases) targets the eradication of a wide variety of diseases, many of which are not prevalent across the globe. As a result, statisticians have selected two statistical indicators, targeting HIV and tuberculosis, as the appropriate global indicators. So not all elements of the target are addressed and thus some elements of the target must be ignored and remain unquantified. While this might make sense from a global perspective, it may not necessarily make sense from a regional or national viewpoint. For example, the control of dengue fever is not a big issue globally, but is very important in South-East Asia. Not surprisingly, when the dust settled, and the indicators had been selected, researchers criticised the indicators for being reductionist (Mair et al. 2018).

Other important decisions are still to be taken. For example, how will changes in the composition of groups be dealt with. Over the course of fifteen years, several Least Developed Countries (LDCs) are likely to graduate from that status. According to the UN Department of Economic and Social Affairs (2019), Vanuatu is expected to graduate in 2020, Angola in 2021, Bhutan in 2023 and São Tomé and Príncipe, the Solomon Islands and perhaps Bangladesh in 2024. What are the implications of this for time-series analyses? Twenty-four of the 169 targets explicitly mention LDCs. When we target an annual growth in GDP of 7% in the LDCs (Target 8.1), a doubling in the share of employment in industry for LDCs (Target 9.2), or a doubling of LDCs' share of global exports (Target 17.11), which LDC group are we referring to? Will rates of change be calculated using the original composition in 2015 or the group as it will be composed in 2030? Or will both series be presented side by side? A relatively straightforward decision, but one where the choice will, most likely, lead to quite different results and may open considerable room for the interpretation of success.

#### 4. Some Tensions in Measuring the SDGs

A surprising discovery emerged during the preparatory work to develop the SDG GIF; many Member States appeared not to fully understand the distinction between national and international official statistics and the significance or purpose of having both. This misunderstanding extended beyond political circles and included also representatives from national statistical offices (NSOs). The SDGs brought this distinction into sharp focus. Confusion around this issue, and subsequent tensions became most acute during the discussions on formulating the '*Guidelines on Data Flows and Global Data Reporting for Sustainable Development Goals*' document (IAEG-SDG 2018b).

Arguably, many of these tensions could have been avoided, if early in the process, the exact scope and purpose of the SDG GIF had been communicated clearly to Member States, and the distinction between global and national indicators had been made clear. Kapto (2019, 135) summarised it well, saying "A tense debate is taking place on data flows from national to regional to global levels, and on custodian agencies' role in harmonising national data for global comparability, as countries assert their sovereignty over national data." The insistence by some Member States that official country data should be prioritised may ultimately be counter-productive given the paucity of data available in many developing countries, resulting in many SDG indicators remaining unpopulated. It may also, inadvertently, undermine the role of international organisations (IOs) that play an important role in compiling harmonised official international statistics, which often involves amending or imputing national data. The 'country first' approach, while to some extent understandable, is nevertheless somewhat incongruous with statements the same countries make vis-à-vis the importance of harnessing the data revolution or using big data and geo-spatial information.

Nevertheless, countries anxious to keep control over messaging are determined that only official national data are used to populate the SDG indicators. Apart from communication control, there are, of course, some legitimate reasons why: (1) national data may be superior, from the perspective of policy formation, as they can be integrated with other national data to present a coherent story; (2) as already noted, the SDG process has expanded the frontiers of official statistics and NSOs may wish to retain control so that they can develop expertise in new statistical domains; (3) NSOs get frustrated when they find results they don't recognise in international databases – often the reasons for the differences are legitimate, but have not been communicated to countries (of course, in some cases, the results were communicated, but some countries didn't pay attention).

There are, however, some circumstances where the 'country first' approach may not necessarily be the best approach. Targets, such as 16.5 (Substantially reduce corruption and bribery in all their forms) or 16.6 (Develop effective, accountable and transparent institutions at all levels), which deal with corruption, bribery and the accountability of institutions, provide perfect examples of why it might make sense to use external or unofficial data, as official data may not exist or may not be sufficiently trustworthy to provide an independent, impartial picture of such sensitive matters. Another exception might be where a single source could provide better-quality and globally more consistent data than the amalgamation of multiple individual country data sets. This might be applicable to targets such as 15.1 that deal with forest, drylands, wetlands and mountain regions governed by international agreements. Arguably, superior quality and internationally comparable data could be derived from satellite imagery.

Using alternative sources to compile official national statistics might also be reasonable where problems with data exist. Problems with data could mean anything from errors or inaccuracies, non-adherence to international standards, incompleteness or data gaps, inconsistencies over time, or imbalances. A good example of where this might arise is the asymmetries that frequently exist between bilateral trade data sets. From a global perspective, unbalanced trade data are not especially useful, and so steps are taken to remove these asymmetries. However, this may lead to a mismatch between official national statistics and official international statistics. This issue is not unique to international trade, 'problems' with national data exist across a range of statistical domains. For the moment, the challenge of how to balance the needs of national and global interests remains unresolved. However, the discussion should not be characterised as national versus international official statistics, but rather how best to integrate and use different statistics to deliver on requirements.

Despite the best efforts of NSOs and IOs, internationally comparable data remains a challenge. The SDG process has exacerbated this challenge, as many of the targets, and consequent indicators, fall well outside the scope of traditional official statistics and thus, are not guided by agreed international measurement standards. Even for those indicators that do fall within the scope of traditional official statistics, there will be a wide variety in general quality and adherence to international standards across countries.

The goals and targets of the 2030 Agenda are underpinned by the ambition that 'no one gets left behind' (UN 2015a). This ambition was translated for statisticians by Mogens Lykketoft, President of the seventieth session of the UN General Assembly, as 'leaving no one uncounted' (Lebada 2016). In principle, this is fine, but such a literal translation does not make much sense from a statistical perspective. The purpose of official statistics, with a few exceptions, such as population censuses, is not to account for every single person but rather to provide general aggregate, anonymised information on population cohorts of interest. This is a fundamental difference between producing official statistics and audited accounts. Apart from issues of confidentiality, the cost of realising the ambition of 'leaving no one uncounted' would be prohibitive and not financially viable for even the best-resourced and

most efficient statistical systems. The challenge for the global statistical system is how to sufficiently improve the granularity of data in a way that prioritises the measurement of the poorest and most vulnerable, that does not divert scarce resources into generating fruitless levels of disaggregation and yet satisfies the demands of political rhetoric.

#### 5. Classifying the SDG Indicators

The far-reaching ambitions of the 2030 Agenda have led to development targets that are well ahead of available official statistics and statistical concepts. In many cases, appropriate statistical definitions and methodologies do not yet exist from which to generate indicators. To elaborate this problem and facilitate the population of the GIF, the IAEG-SDG (2018a) has classified all SDG indicators into three tiers based on their conceptual development and availability of data. The tiers are:

Tier 1: the indicator is conceptually clear, has an internationally established methodology, standards are available, and data are regularly produced by countries for at least 50% of countries and of the population in every region where the indicator is relevant.

Tier 2: the indicator is conceptually clear, has an internationally established methodology, standards are available, but data are not regularly produced by countries.

Tier 3: no internationally established methodology or standards are yet available for the indicator, but methodology/standards are being (or will be) developed or tested.

In September 2019, the IAEG-SDG reported that 45% of the selected indicators were classified as Tier 1 (see Table 1). Furthermore, they reported that 14% of the indicators remained classified as Tier 3. While Table 1 shows the not inconsiderable improvements in conceptual development and data availability that has been made since 2016, it also highlights the magnitude of the task still facing the global statistical community. The pace of transition of indicators through the tiers to reach Tier 1 is likely to slow, as presumably the low hanging fruit will be picked first. Table 1 suggests this is indeed the case, as the conversion rate to Tier 1 was slower between December 2017–2018 than between December 2016–2017. Between December 2018 and September 2019 only three indicators were converted to Tier 1. A further cautionary footnote should be added. Research undertaken by Dang and Serrjuddin (2019) of the World Bank

| Tier<br>Classification | December<br>2016 |     | December<br>2017 |     | December<br>2018 |     | September<br>2019 |     |
|------------------------|------------------|-----|------------------|-----|------------------|-----|-------------------|-----|
|                        | Number           | %   | Number           | %   | Number           | %   | Number            | %   |
| 1                      | 81               | 35  | 93               | 40  | 101              | 44  | 104               | 45  |
| 2                      | 57               | 25  | 66               | 28  | 84               | 36  | 89                | 38  |
| 3                      | 88               | 38  | 68               | 29  | 41               | 18  | 33                | 14  |
| Multiple               | 4                | 2   | 5                | 2   | 6                | 3   | 6                 | 3   |
| Total                  | 230              | 100 | 232              | 100 | 232              | 100 | 232               | 100 |

Table 1. Number of SDG indicators by tier.

Source: Derived from IAEG-SDG (2019b). https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/

highlights the 'overwhelming challenge with missing data' and suggests that not all Tier 1 indicators are actually populated. They estimate that only 19% of the required GIF data are currently available.

#### 6. The Cost of Measurement

One of the implications of such a broad and ambitious development agenda is the price tag. Estimates vary, but Ambassador Macharia Kamau of Kenya, who co-chaired the SDG intergovernmental consultative process, estimates that implementing the SDG agenda could cost somewhere between USD 3.5 trillion and USD 5 trillion per year (Deen 2015). The Economist (2015a) described their estimate, of between USD 2 trillion and USD 3 trillion per year (or the equivalent of 4% of global GDP), as 'unfeasibly expensive'. The Intergovernmental Committee of Experts on Sustainable Development Financing (2014) estimated the value of investment in infrastructure required to achieve the eradication of poverty alone at between USD 5 trillion and USD 7 trillion annually.

Even for developed countries with relatively advanced and sophisticated statistical systems, the demands arising from the SDG monitoring framework are immense. When one considers that in 2019, only 45% of the proposed 232 indicators were classified as Tier 1 (see Table 1), the extent of the problem becomes clear. PARIS21 (2015, 11) has estimated that 'funding for statistics needs to be increased from current commitments of between USD 300 million and USD 500 million to between USD 1 billion and USD 1.25 billion by 2020'. The Global Partnership for Sustainable Development Data estimates that around USD 650 million per year is needed to collect data, of which only USD 250 million is currently funded (Runde 2017). Irrespective of which estimate is used, these sums clearly exceed existing funding (UNCTAD 2016). While clearly the bulk of these resources will be required to improve statistical capacity in developing countries, it is evident that resources will be required in the developed world too in order to deliver on the promises made by national governments.

Although statistics account for only 0.3% of official development assistance (ODA) (PARIS21 2017), USD 541 million is not a trivial amount of money and the thought of paying more doesn't appear to excite many donors. In this context, Slotin (2018) asks a relevant question 'if development data is so powerful, why does no one want to pay for it?' Assuming the answer is a poor understanding of the contribution of official statistics as a public good to democracy, commerce and social wellbeing, official statisticians have set out to show it's a price worth paying, arguing 'if you think statistics are expensive – try ignorance'. To try to justify this claim and the costs of measurement, a nascent industry is now emerging, where statisticians are trying to estimate the benefits of official statistics. Chui et al. (2013) found that open data globally could potentially unlock between USD 3.2 trillion and USD 5.4 trillion in economic value per year. UNECE (2019) has compiled a report on the various methods used and cites a variety of estimates. The numbers are seductive, but it doesn't change the fact that, as yet, NSOs are not getting additional funding to compile SDG indicators.

## 7. Some Unanticipated Consequences

According to the American cultural and intellectual historian T.J. Jackson Lears, 'All history is the law of unintended consequences' (Cohen 2013). It should not be surprising

then that a development plan as broad as the 2030 Agenda and the implementation of the SDG GIF should throw up a few surprises. This section of the article examines what some of these surprises are and what the consequences might be.

The delegation of the selection and measurement of the statistical indicators to the UN Statistical Commission was a major triumph for official statistics. It was an explicit recognition of the need for apolitical, independent and impartial official statistics to measure progress and the separation of function between statistical compilers and statistical consumers. It also responded to the views expressed by many that the 2030 Agenda needed an effective performance system with clear metrics measuring progress towards each goal (Warren 2015; Costanza et al. 2016; Jacob 2017).

However, SDG targets are not 'targets' in the normal sense of the word – they are, for the most part, not clear time-delimited objectives but rather general, often complex, aspirations that leave generous space for interpretation. Furthermore, they incorporate all the unresolved issues left over from the negotiation phase. Thus, in handing over the measurement task to the statistical community, the interpretation of the targets was effectively delegated too. Many heads of state and policy mandarins might be surprised by, or even contest, this statement. However, it was statisticians who selected the indicators that specifically defined what the 2030 Agenda text actually meant. This is an important point because the SDG indicators do not simply measure the 2030 Agenda, they define it. As noted above, the composition of the dashboard selected to measure Target 17.13 will effectively determine whether the 2030 Agenda adopts an orthodox or heterodox view of the global economy. Equally, how statisticians interpret the word 'illicit' when designing indicator Target 16.4 (Total value of inward and outward illicit financial flows in current USD) will determine whether corporate profit shifting is included or not. In making these decisions, statisticians are effectively determining what the SDG targets mean. The indicators selected are the performance metrics for the 2030 Agenda and thus will have direct consequences for whether the 2030 Agenda is judged a success or a failure.

With 169 SDG targets, many of which are multidimensional, there was an understandable fear of indicator proliferation. After all, the MDGs had only 21 targets but 60 indicators, a ratio of 3:1. The prospect of 500 plus indicators was not attractive to politicians. Hence the limit of one indicator per target. Yet there are consequences to measuring a multifaceted target with a single indicator. The first and most obvious being that, unless a composite indicator or a multidimensional dashboard can be designed, several elements of the target will be sacrificed. This may be appropriate if the target dimensions are all somehow related. However, in several cases, the dimensions included in the targets do not appear to be related at all, in which case, any single indicator will be problematic. A single indicator thus introduces the risk that unmeasured aspects of a target will be ignored, and interconnections between different elements of the target (or other targets) will remain unquantified. An obvious worry from a policy perspective are the implications for consequent behaviour and the risk that *only* what gets measured gets done.

Another concern is how to extrapolate from narrow indicators to broad targets. For example, Target 17.4 (Assist developing countries in attaining long-term debt sustainability through coordinated policies aimed at fostering debt financing, debt relief and debt restructuring, as appropriate, and address the external debt of highly indebted poor countries (HIPCs) to reduce debt distress) is a complex, multidimensional target represented by a single indicator. In this case, a variety of complex issues, such as: long-term debt sustainability; debt financing; debt relief; debt restructuring; and external debt have all been shoe-horned into a single indicator. Furthermore, indicator 17.4.1 doesn't really address any of the target elements directly or adequately, raising questions as to how progress towards Target 17.4 should be interpreted.

Very few of the SDG indicators are bespoke indicators that fit the specifications of the target exactly. Furthermore, very few were deliberately designed for the purpose for which they are now being used; most are to some extent or other, recycled, proxies. This will matter when the trends and patterns identified by the indicator are extrapolated and applied to all elements of the broader target. It is important to understand what that original purpose of the indicator was, so that its appropriateness as an SDG indicator can be assessed. The small print (otherwise known as metadata) will be very important when analysing the SDG results.

### 8. Discussion

The 2030 Agenda may have a profound influence on the shape and organisation of official statistics in the future. As noted above, many policy discussions are running far ahead of available statistics, and so the SDGs are likely to be the driving force, or *raison d'être*, for many statistical advances in the coming years, both in terms of statistical concepts and methodology but also in terms of statistical organisation and the use of new data sources. It is very important that all national statistical systems (NSSs) engage actively in these discussions. As Harari (2018, ix) notes 'history gives no discounts.' Countries that do not engage will not be exempt from the consequences. This may have three unexpected outcomes:

- Statistical organisation the demand for new statistics may inadvertently open the door to the outsourcing or privatisation of official statistics if the existing system fails to deliver on the huge expectations that appear to exist. There is a risk that if the UN statistical system cannot fill the vacuum created by the Tier II and III indicators, then someone else will. This is not necessarily a bad thing; there are many who argue that a more inclusive approach (MacFeely and Nastav 2019) or incorporating citizen science (Fritz et al. 2019) might benefit the SDG measurement process.
- 2. Big Data there is growing pressure on NSOs to try to harness big data to compile statistics. This is, in of itself not problematic, although expectations should be realistic (see MacFeely 2019), but it may distract from developing administrative data sources, which arguably will be more useful (in the short to medium term at least). The real challenge, and optimal objective, is for NSOs to find a way to integrate multiple data sources, whether traditional or new, to develop efficient national statistical and information systems (UNECE 2016; MacFeely and Barnat 2017).
- 3. Reputational risk given the very short timeframe in which the GIF was developed without any appreciable additional resources, a lot has been achieved. Nevertheless, the SDG GIF may still disappoint the high expectations, and this in turn may undermine the UN Statistical Commission (MacFeely and Nastav 2019). On the

other hand, it also offers an opportunity to re-engage with policy makers and discuss the importance of official statistics.

Thus, for a variety of reasons the SDG GIF is likely to have lasting implications beyond the 2030 Agenda. NSOs need to reflect carefully on these issues. A particular challenge posed by the 2030 Agenda for statisticians is that some of the SDG targets deal with phenomena that arguably cannot be measured comprehensively, if at all. Cited above, Target 4.7 is a good example. Target 17.16 (Enhance the Global Partnership for Sustainable Development, complemented by multi-stakeholder partnerships that mobilize and share knowledge, expertise, technology and financial resources, to support the achievement of the Sustainable Development Goals in all countries, in particular developing countries) is another. This is not a criticism of the target or the aspirations contained therein, simply that some issues are by their nature nebulous and defy robust quantitative measurement. As quantification and metrics have irrevocably become part of society's zeitgeist, no one is questioning whether this approach is sensible or achievable – it is now a commonly held view that everything can and must be measured. In an era of governance by numbers, the management clichés of 'measure what you treasure' or 'what gets measured gets done' rule supreme. However, as Muller (2018, 8) points out, 'measurement may become counterproductive when it tries to measure the unmeasurable and quantify the unquantifiable.' While no one can credibly challenge the logic of evidence-informed decision making, arguably statisticians could also play an important role in advising what can and cannot be sensibly measured. If the SDG GIF is to be useful, it is essential that users understand the limitations of these types of performance indicators. This will be especially important for donors who could conceivably make funding decisions conditional on these indicators.

The SDG indicators have hijacked, to some extent, the discussion on which statistics and data are required to support sustainable development. What has often been lost in the debate thus far is that the SDG indicators are only performance metrics – they will tell us whether a target is being achieved or not. This focus on indicators risks relegating statistics to the downstream role of monitoring and evaluation. A key role of statistics should be to inform policy decisions – this upstream or diagnostic role seems to have been, to some extent, lost in discussion. There has been relatively little debate on what additional data are required to inform and design integrated policies in order to implement actions to achieve the SDG targets. It is important that the data and statistics required to undertake risk assessments, formulate policy or design early warning systems are not forgotten during the discussions on resource mobilisation or capacity development.

#### 9. Conclusions and Recommendations

The 2030 Agenda represents the first ever democratically forged agreement on development and will guide global development for the next ten to fifteen years. That agenda will also guide many new statistical developments and will be the driving force behind the breaking of new statistical ground.

The UN Statistical Commission and the IAEG-SDG has made tremendous progress. Despite many constraints, the SDG GIF was assembled in record time. Nevertheless, critics of the SDG indicators have criticised them for being reductionist and of watering down the ambition of the goals and targets (Fukuda-Parr and McNeill 2019; Engle Merry 2019; Razavi 2019). Yet statistics are by definition reductionist. The question is whether in the unavoidable distillation process, the essence of the target has been faithfully captured or not. There is no question that some indicators have missed their targets and others are probably watered down from the ambitions of the target. This was almost unavoidable, given the complexity of most targets and the requirement to have only one indicator per target.

While many of the criticisms hold some water, the SDG indicator process has arguably focused more attention on global official statistics than any other UN programme. Many politicians and diplomats are now beginning to understand some of the challenges associated with consistently measuring development issues. While there are issues surrounding some indicators, they hopefully offer, at least some common ground to progress policy discussions. The SDG process also offers an opportunity for statisticians to engage and reflect with data users on the future direction of official statistics. In this context, four key strategic issues are outlined below. Without question, others could be added, but the issues highlighted here must be central to any serious discussion regarding the strategic role of NSSs and the international statistical system in the future.

- 1. NSOs could consider broadening their mandate to include the homologation of statistics created by third parties. Such a move would probably be welcomed by nongovernment organisations, civil society and academia - perhaps even the private sector, and would certainly be in keeping with the inclusive spirit of the 2030 Agenda. It would also help to maintain quality control and promote sound methodologies, transparency and openness of data (Cervera et al. 2014; Landefeld 2014; Kitchin 2015; MacFeely 2016; Hammer et al. 2017). The challenge, of course, will be for NSOs to acquire the expertise to conduct thorough and professional homologation as the frontiers of official statistics are broadening so quickly. At the global level, the United Nations could be more proactive and introduce an accreditation system (with uniform standards) that would allow unofficial compilers of statistical indicators to be accredited as 'official' for the purposes of populating the SDG GIF. One could envisage, for example, the IAEG-SDG or a similar body with the authority and competence to certify statistics as 'fit for purpose' reviewing unofficial statistics to see whether they can be certified as 'official' for the purposes of populating the SDG global monitoring framework. Without going into detail, this approach would be suitable for Tier 3 or Tier 2 indicators that otherwise run the risk of remaining unpopulated. By encouraging more active participation in the measurement, such an approach might help to domesticate the 2030 Agenda and reduce the costs of populating the SDG GIF. A detailed discussion of this idea can be found in MacFeely and Nastav (2019).
- 2. Newly emerging globalised digital data also offer exciting possibilities and opportunities to reconsider the national production models currently employed by NSOs and NSSs. Switching from a national to a collaborative international production model might make sense from an efficiency or international comparability perspective, but it would be a dramatic change in approach, and

possibly a bridge too far for many NSOs and governments. Globalised data are already presenting challenges as they defy national sovereignty, putting the owners and the data themselves beyond the reach of national legal systems. Governments, already struggling to enforce national laws and protect citizens, now recognise this as an important policy issue (Casalini and López González 2019; UNCTAD 2019). Nevertheless, global digital data offer opportunities to consider centralising some statistical production in a single centre, offering real international comparability, rather than replicating production many times over in individual countries. Obviously, this would not work for all domains, as issues like scale matter, as does integration with other local or national data sets. As noted above, the challenge is how to efficiently integrate the variety of data sources now available in a way that allows statistics to meet both local and global policy needs.

- 3. The 2030 Agenda has provided yet further justification, if further justification were needed, that countries should develop their NSSs and, to develop their NSSs and put in place a national data infrastructure (UNCTAD 2016; UNESCAP 2019). Possibly because most official statistics and disseminated administrative data are viewed as a public good, their value is not well understood or fully appreciated. Politicians do not always understand the concept of soft or nonphysical infrastructure and so may find this argument nebulous. The United Nations should take this opportunity to explain to countries that in an information age, data are an economic resource and a strategic asset, and that administrative data are not an unfortunate cost but rather a valuable national asset. Governments should also be helped to understand that data infrastructure is every bit as important as broadband or pipelines. Furthermore, the UN should emphasise that a national data infrastructure will not happen by itself, but with careful architectural design, can contribute to public sector efficiency, as well as better statistics to support public policy design and evaluation (MacFeely and Dunne 2014).
- 4. At the time of writing, we are almost one-third of way through the 2030 Agenda. Preparations for the post-2030 Agenda will mostly likely begin in 2028 – only seven years from now. The statistical community should prepare well in advance of the post-2030 debates. It is important that statisticians reflect on the SDG and IAEG-SDG processes and learn lessons. What worked and what didn't? How do we avoid making the same mistakes again? As a statistical community, what would we like to see changed in the follow-up programme?

This article has outlined some of the measurement challenges, tensions, unexpected consequences and strategic issues for statistics emerging from the 2030 Agenda. Again, it should be stressed that unanticipated consequences are not necessarily a bad thing. The 2030 Agenda may have inadvertently opened up new and unexpected opportunities to reimagine the traditional role of official statistics – to engage in new partnerships and build wider data ecosystems, and to develop new statistical concepts and methodologies. The ambition of the 2030 Agenda arguably provides an open door to consider bolder solutions. On the other hand, some emerging, and perhaps unexpected, clouds can also be seen on the horizon. It is not clear whether the statistical community has yet given sufficient thought to these.

#### 10. References

- Annan, K. 2000. We the peoples: The role of the United Nations in the 21st century. Available at: http://www.un.org/en/events/pastevents/pdfs/We\_The\_Peoples.pdf (accessed February 2017).
- Casalini, F. and J. López González. 2019. *Trade and Cross-Border Data Flows*. OECD Trade Policy Papers, No. 220, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/b2023a47-en.
- Cervera, J.L., P. Votta, D. Fazio, M. Scannapieco, R. Brennenraedts, and T. van der Vorst. 2014. *Big Data in Official Statistics*. Eurostat ESS Big Data Event, Rome 2014 – Technical Event Report. Available at: https://ec.europa.eu/eurostat/cros/system/files/ Big%20Data%20Event%202014%20-%20Technical%20Final%20Report%20-final V01\_0.pdf (accessed January 2018).
- Chui, M.J., D. Farrell, S. van Kuiken, P. Groves, and E.A. Doshi. 2013. Open Data: Unlocking innovation and performance with liquid information. McKinsey Digital, McKinsey Global Institute. Available at: https://www.mckinsey.com/business-functions/ mckinsey-digital/our-insights/open-data-unlocking-innovation-and-performance-withliquid-information (accessed November 2019).
- Cohen, B.R. 2013. "The Confidence Economy: An Interview with T. J. Jackson Lear." Available at: https://www.publicbooks.org/the-confidence-economy-an-interview-with-t-j-jackson-lears/ (accessed June 2019).
- Costanza, R., L. Daly, L. Fioramonti, E. Giovannini, I. Kubiszewski, L.F. Mortensen, K.E. Pickett, V.K. Ragnarsdottir, R. de Vogli, and R. Wilkinson. 2016. "Modelling and measuring sustainable wellbeing in connection with the UN Sustainable Development Goals." *Ecological Economics* 130: 350–355. DOI: https://doi.org/10.1016/j.ecolecon.2016.07.009.
- Dang, H.H. and U. Serajuddin. 2019. *Tracking the Sustainable Development Goals: Emerging Measurement Challenges and Further Reflections*. World Bank Policy Research Working Paper, No. 8843. 7 May, 2019. Available at: https://papers.ssrn.com/ sol3/papers.cfm?abstract\_id=3383814 (accessed May 2019).
- Deen, T. 2015. "UN targets trillions of dollars to implement sustainable development agenda." Global Policy Forum. Available at: https://www.globalpolicy.org/component/ content/article/271-general/52800-un-targets-trillions-of-dollars-to-implement-sustainable-development-agenda.html (accessed December 2019).
- Engle Merry, S. 2019. "The Sustainable Development Goals Confront the Infrastructure of Measurement." *Global Policy* 10(1): 145–146. DOI: https://doi.org/10.1111/ 1758-5899.12606.
- Fritz, S., L. See, T. Carlson, M. Haklay, J.L. Oliver, D. Fraisl, R. Mondardini, M. Brocklehurst, L.A. Shanley, S. Schade, U. When, T. Abrate, J. Anstee, S. Arnold, M. Billot, J. Campbell, J. Espey, M. Gold, G. Hager, S. He, L. Hepburn, A. Hsu, D. Long, J. Masó, I. McCallum, M. Muniafu, I. Moorthy, M. Obersteiner, A.J. Parker, M. Weissplug, and S. West. 2019. "Citizen science and the United Nations Sustainable Development Goals." *Nature Sustainability* 2: 922–930. DOI: https://doi.org/10.1038/s41893-019-0390-3.
- Fukuda-Parr, S. and D. McNeill. 2019. "Knowledge and Politics in Setting and Measuring the SDGs: Introduction to Special Issue." *Global Policy* 10(1): 5–15. DOI: https://doi.org/10.1111/1758-5899.12604.

- Hammer, C.L., D.C. Kostroch, G. Quiros, and STA Internal Group. 2017. *Big Data: Potential, Challenges, and Statistical Implications.* IMF Staff Discussion Note, SDN/17/06, September 2017. Available at: http://www.imf.org/en/Publications/SPROLLs/ Staff-Discussion-Notes (accessed January 2018).
- Harari, Y.N. 2018. 21 Lessons for the 21st Century. London: Jonathan Cape.
- Hulme, D. 2009. The Millennium Development Goals (MDGs): A short history of the world's biggest promise. BWPI Working Paper 100, Brooks World Poverty Institute, University of Manchester. Available at: https://www.unidev.info/Portals/0/pdf/bwpi-wp-10009.pdf (accessed June 2018).
- IAEG-SDG (Inter-Agency and Expert Group on Sustainable Development Goal Indicators). 2018a. *Tier Classification for Global SDG Indicators*. Available at: https://unstats.un.org/sdgs/iaeg-sdgs/ (accessed June 2018).
- IAEG-SDG (Inter-Agency and Expert Group on Sustainable Development Goal Indicators). 2018b. *Guidelines on Data Flows and Global Data Reporting for Sustainable Development Goals*. Background Document to Item 3 (a) of the Forty-ninth session of the United Nations Statistical Commission. Available at: https://unstats.un. org/unsd/statcom/49th-session/documents/BG-Item-3a-IAEG-SDGs-DataFlowsGuidelines-E.pdf (accessed December 2018).
- IAEG-SDG (Inter-Agency and Expert Group on Sustainable Development Goal Indicators). 2019. Data Disaggregation and SDG Indicators: Policy Priorities and Current and Future Disaggregation Plans – Prepared by the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs). Statistical Commission, Fiftieth session. Available at: https://unstats.un.org/unsd/statcom/50th-session/documents/ BG-Item3a-Data-Disaggregation-E.pdf (accessed July 2019).
- IAEG-SDG (Inter-Agency and Expert Group on Sustainable Development Goal Indicators). 2019b. *Tier Classification Review September 2019*. Available at: https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/ (accessed November 2019).
- Intergovernmental Committee of Experts on Sustainable Development Financing. 2014. *Report of the Intergovernmental Committee of Experts on Sustainable Development Financing – Final Draft, 8 August 2014.* Available at: https://sustainabledevelopment. un.org/content/documents/4588FINAL%20REPORT%20ICESDF.pdf (accessed February 2017).
- Jacob, A. 2017. "Mind the Gap: Analyzing the Impact of Data Gap in Millennium Development Goals' (MDGs) Indicators on the Progress toward MDGs." *World Development* 93: 260–278. DOI: https://doi.org/10.1016/j.worlddev.2016.12.016.
- Jütting, J. 2016. *Capacity building, yes but how to do it?* Available at: http://undataforum.org/WorldDataForum/capacity-building-yes-but-howto-do-it/ (accessed November 2017).
- Kapto, S. 2019. "Layers of Politics and Power Struggles in the SDG Indicators Process." *Global Policy* 10(1): 134–136. DOI: https://doi.org/10.1111/1758-5899.12630.
- Kitchin, R. 2015. "The opportunities, challenges and risks of big data for official statistics." *Statistical Journal of the International Association of Official Statistics* 31(3): 471–481. DOI: https://doi.org/10.3233/SJI-150906.

- Landefeld, S. 2014. "Uses of Big Data for Official Statistics: Privacy, Incentives, Statistical Challenges, and Other Issues." Discussion paper presented at the United Nations Global Working Group on Big Data for Official Statistics, Beijing, China, 31 October 2014. Available at: https://unstats.un.org/unsd/trade/events/2014/beijing/Steve%20Landefeld%20-%20Uses%20of%20Big%20Data%20for%20official%20statistics.pdf (accessed January 2018).
- Lebada, A.M. 2016. *Member states, statisticians address SDG monitoring requirements.* Available at: http://sdg.iisd.org/news/member-states-statisticians-address-sdg-monitoring-requirements/ (accessed December 2019).
- MacFeely, S. 2016. "The continuing evolution of official statistics: Some challenges and opportunities." *Journal of Official Statistics* 32(4): 789–810. DOI: https://doi.org/ 10.1515/jos-2016-0041.
- MacFeely, S. 2019. "The Big (data) Bang: opportunities and challenges for compiling SDG indicators." *Global Policy* 10(1): 121–133. DOI: https://doi.org/10.1111/1758-5899.12595.
- MacFeely, S. and N. Barnat. 2017. "Statistical capacity building for sustainable development: Developing the fundamental pillars necessary for modern national statistical systems." *Journal of the International Association of Official Statistics* 33(4): 895–909. DOI: https://doi.org/10.3233/SJI-160331.
- MacFeely, S. and J. Dunne. 2014. "Joining up public service information: The rationale for a national data infrastructure." *Administration* 61(4): 93–107. Available at: https://cora.ucc.ie/handle/10468/9513?show=full.
- MacFeely, S. and B. Nastav. 2019. "You say you want a [data] revolution. A proposal to use unofficial statistics for the SDG Global Indicator Framework." *Journal of the International Association of Official Statistics* 35(3) (forthcoming).
- Mair, S., A. Jones, J. Ward, I. Christie, A. Druckman, and F. Lyon. 2018. A Critical Review of the Role of Indicators in Implementing the Sustainable Development Goals.
  In: Leal Filho, W, (ed.) Handbook of Sustainability Science and Research. Manchester, UK. Springer, 41–56.
- Muller, J.Z. 2018. The Tyranny of Metrics. USA: Princeton University Press, Princeton.
- PARIS21. 2015. A road map for a country-led data revolution. Available at: http://www. oecd-ilibrary.org/docserver/download/4315051e.pdf?expires=1457406953&id=id &accname=guest&checksum=6B4747834B1E459F5E186E65EE1034B5 (accessed January 2017).
- PARIS21. 2017. Partner Report on Support to Statistics Press 2017. Available at: https://paris21.org/sites/default/files/2017-10/PRESS2017\_web2.pdf (accessed November 2017).
- Razavi, S. 2019. "Indicators as Substitute for Policy Contestation and Accountability? Some Reflections on the 2030 Agenda from the Perspective of Gender Equality and Women's Rights." *Global Policy* 10(1): 149–152. DOI: https://doi.org/10.1111/1758-5899.12633.
- Runde, D. 2017. The Data Revolution in Developing Countries Has a Long Way to Go Forbes. 25 February, 2017. Available at: https://www.forbes.com/sites/danielrunde/ 2017/02/25/the-data-revolution-in-developing-countries-has-a-long-way-to-go/# 620717201bfc (accessed September 2018).

- Slotin, J. 2018. *What Do We Know About the Value of Data?* Global Partnership for Sustainable Development Data. Available at: http://www.data4sdgs.org/sites/ default/files/services\_files/Value%20of%20Data%20Report\_Final\_compressed\_0.pdf (accessed November 2019).
- The Economist. 2015a. "The 169 commandments." 28 March 2015. *The Economist*. Available at: https://www.economist.com/leaders/2015/03/26/the-169-commandments (accessed December 2019).
- The Economist. 2015b. "Assessing development goals: The good, the bad and the hideous." 26 March 2015. *The Economist*. Available at: https://www.economist.com/international/ 2015/03/26/the-good-the-bad-and-the-hideous (accessed December 2019).
- UN (United Nations). 2013. Open Working Group proposal for Sustainable Development Goals. Available at: https://sustainabledevelopment.un.org/content/documents/1579SDGs% 20Proposal.pdf (accessed February 2017).
- UN (United Nations). 2015a. Transforming our world: the 2030 Agenda for Sustainable Development Resolution 70/1 adopted by the General Assembly on 25 September 2015. Available at: http://www.un.org/ga/search/view\_doc.asp?symbol=A/RES/70/1& Lang=E (accessed February 2017).
- UN (United Nations). 2015b. *Agenda, 2030 "to-do list for people and planet*", Secretary-General Tells World Leaders Ahead of Adoption [press release]. Available at: http://www.un.org/press/en/2015/sgsm17111.doc.htm (accessed January 2015).
- UN (United Nations). 2000. United Nations Millennium Declaration. Resolution 55/2 adopted by the General Assembly. Available at: http://www.un.org/en/ga/search/view\_doc.asp?symbol=A/RES/55/2 (accessed February 2017).
- UNCTAD (United Nations Conference on Trade and Development). 2016. *Development and globalisation: Facts and Figures 2016*. Available at: http://stats.unctad.org/ Dgff2016/index.html (accessed October 2016).
- UNCTAD (United Nations Conference on Trade and Development). 2019. *Digital Economy Report 2019 – Value Creation and Capture: Implications for Developing Countries*. Available at: https://unctad.org/en/pages/PublicationWebflyer.aspx?publicationid=2466 (accessed November 2019).
- UN (United Nations) Department of Economic and Social Affairs. 2019. Graduation from the LDC category: Timeline Country Graduations. Available at: https://www.u-n.org/development/desa/dpad/least-developed-country-category/ldc-graduation.html (accessed December 2019).
- UNECE (United Nations Economic Commission for Europe). 2016. Outcomes of the UNECE Project on Using Big Data for Official Statistics. Available at: https://statswiki.unece.org/display/bigdata/Big+Data+in+Official+Statistics (accessed February 2018).
- UNECE (United Nations Economic Commission for Europe). 2019. *Recommendations for Promoting, Measuring and Communicating the Value of Official Statistics*. Available at: http://www.unece.org/index.php?id=51139 (accessed November 2019).
- UN (United Nations) Environmental Programme. 2015. *Clarifying terms in the SDGs: Representing the meaning behind the terminology*. Available at: http://unstats.un.org/ sdgs/files/meetings/iaeg-sdgs-meeting-02/Statements/UNEP%20-%20Clarifying% 20terms%20in%20the%20SDGs.pdf (accessed January 2017).

- UNESCAP (United Nations Economic and Social Commission for Asia and the Pacific. 2019. *Integrated Statistics: A journey worthwhile*. Stats Brief. July 2019, Issue No. 19. Available at: https://www.unescap.org/sites/default/files/Stats\_Brief\_Issue19\_Jul 2019\_Integrated\_Statistics.pdf (accessed November 2019).
- UN (United Nations) General Assembly. 2017. Resolution adopted by the General Assembly on 6 July 2017 Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development. Seventy-first session. A/RES/71/313. Available at: https://undocs.org/A/RES/71/313 (accessed June 2019).
- UN (United Nations) Statistical Commission. 2017. *Report on the forty-eighth session* (7–10 March 2017). Economic and Social Council, Official Records 2017 Supplement No. 4, E/2017/24-E/CN.3/2017/35. Available at: https://undocs.org/pdf? symbol=en/E/2017/24.
- Warren, S. 2015. Analyzing the #SDGs on Twitter: The social media response to the Sustainable Development Goals. Available at: https://www.owler.com/reports/dalberg-global/dalberg-global-blog-analyzing-the-sdgs-on-twitter/1449841459497 (accessed December 2019).

Received June 2019 Revised November 2019 Accepted December 2019