Standards as the backbone for official statistics, how well do they fit within national and international systems?[1,2] or Does the obsession with cross-national comparisons blind us to the weakly implemented standards behind them?

Pieter Everaers
Editor in Chief, Statistical Journal of the IAOS
E-mail: pevssjiaos@gmail.com

Abstract. Standards, guidelines, and recommendations[3] in statistics are undoubtedly important building blocks of modern global official statistics. They are main achievements of the harmonization work of the global community of official statisticians and are the backbone of the system, essential for maintaining high quality by using well-developed and documented methodologies and procedures, and for facilitating cross-national and cross-regional comparison of scores on indicators and trends. However, beyond such positive assessed achievements, the development, implementation, and use of 'standards' can also be characterized by a variety of negatively assessed characteristics. For example, in the development of standards, not every situation of a country or region can be or is sufficiently weighted or involved, countries and regions lack resources for full implementation, and the use of the standards according to the 'required operational level' is insufficient or the 'standards' do not fit the socio-economic, cultural, or political situation in a country resulting in a non-valid picture and unfit for use in policymaking.

Therefore, though the 'standards' are the backbone of modern official statistics it is valid to question if the frequent use of cross-national comparisons unequally dismisses the cultural specificities of a country or region. It is also relevant to ask if the standards that are used to produce the indicators for cross-national comparisons are sufficiently implemented to allow for valid comparisons. And in general, one might question if there is a misfit between the emphasis on and practice of cross-national comparisons by international organizations and the attention to the level and awareness of the implementation of the standards used to produce the indicators on the country level.

Keywords: Harmonization, statistical standards, guidelines, recommendations

---

[1]This manuscript has two origins. On the one hand, the discussions and presentations at the Special Invited Paper Session at the 2021 ISI world Statistics Conference [1–5] and the author's own experiences in the statistical capacity building led to a reflection on the use of statistical standards in less-developed statistical systems. On the other hand, based on an email exchange with Denise Lievesley, and later joined by Jenny Church, Phil Crook, and Steve MacFeely, the issue of cultural specificity and the validity of cross-national comparisons based on so-called harmonized indicators was introduced.

[2]The intention is to further discuss these questions in expert group

meetings and to develop based on the outcomes of these discussions an extended position paper to be published for wider use and referencing. The 12th discussion on the SJIAOS discussion platform is part of the intended wider discussion on this topic.

[3]In this manuscript internationally harmonized standards, guidelines, and recommendations in official statistics are jointly labeled as 'standards', even though their objectives, level of detail, and reach can be different.

1. Introduction

Standards, guidelines, and recommendations in statistics are undoubtedly important building blocks of modern global official statistics. They are the backbone of the statistical system, essential for maintaining high quality by using well-developed and documented harmonized methodologies and procedures, and for facilitating cross-national and cross-regional comparison of scores on indicators and trends. Standardization is considered positive from these perspectives of improving comparability and allowing regional or global monitoring as well as having procedures that when applied and well-documented witness the quality of the process and consequently the results.

However, beyond such positive achievements, the development, implementation, and use of 'harmonized standards' can also be characterized by a variety of negatively assessed characteristics or challenges. Examples of such characteristics are the following:

A bias in the development of the standards, focussing on a specific set of (more developed) countries or regions, supported by experts that are on different levels of expertise and on the knowledge that is not equally spread.

A different politically inspired interest of countries and regions for a certain standard, or the misfit of a classification developed for global statistical purposes with local or regional desired classifications.

The resources both financial and human needed to implement standards, the knowledge and basic data needed as well as the political will are also influencing the level of implementation and operationalization of standards in countries and regions.

A straw poll of statisticians from some 15 developing statistical systems that are involved in producing indicators used for making cross-national comparisons, supported the idea that comparative statistics are only partially valid due to these characteristics.

Before detailing in paragraphs 4 to 8 these challenges for cross-national comparisons, the role of internationally agreed standards will be described (par 2), as well as a brief history, will be given of harmonized standards in statistics (par 3).

In summary, this manuscript aims to discuss

– If the underlying principles by which internationally agreed on statistical standards and classifications are developed do reflect too heavily the circumstances and preoccupations of high-income countries, thus reducing their relevance to LMICs?

– If the standards that are used to produce the indicators for cross-national comparisons are sufficiently well implemented across the globe to allow for valid international comparisons?
– If the frequent use of and emphasis on cross-national comparisons lead to the cultural specificities of a country or region being given insufficient attention?

2. The role of internationally agreed standards in official statistics

There are two major motivations for the creation of internationally agreed standards for official statistics. The first and most obvious is that there are many reasons why policymakers and commentators, either nationally or internationally based, will wish to compare the situation of different countries or regions (and overtime) using statistics. And for those comparisons to be valid the statistics need to be based on harmonized methods, classifications, nomenclatures, definitions, and so on. So, the objective of these internationally harmonized standards is to be able to collect statistical information on certain phenomena that can be compared without having to reflect on different methodologies, classifications, etc.

The second motivation is that statistical standards are needed for national purposes as well as international ones, and their creation is a very labor-intensive and time-consuming business. The use of well-documented standards is considered an important element for guaranteeing also over time a consistent quality level. Therefore, it makes sense for countries to collaborate rather than each 'reinvents the wheel'.

3. A brief history of standards in statistics[4]

During the last circa 50 years, official statistics have developed from individual national systems with each their specific procedures, techniques and methodologies, to a global statistical system[5] where countries and

---

[4]The Statistical Journal of the IAOS has scheduled for Volume 39 (2023) a Special Issue of the journal on The History of Official Statistics. It is expected that in this issue several manuscripts will deal in detail with the role of (specific) standards in the development of official statistics.

[5]'A 'global statistical system' is used, though, in fact, it is a series of several (party similar) systems, that depending on the statistical domain, more or fewer countries can subscribe to. To some standards, almost all UN member countries subscribe and actively use them as to others fewer countries comply.

regions are using the same set of procedures, techniques and methodologies. Most of these harmonized standards in use in the global statistical system have been developed during the last 20 to 30 years. This development took place thanks to initiatives of individual statisticians and national statistical offices and the enhanced international coordination for example via the United Nations Statistical Commission (UNSC) meetings (discussing the need and requirements) and the international statistical organizations.[6] The statistical systems in those countries where policy-making based on statistical information has a long history, in general also the more advanced countries and regions, were in the lead for expressing the need for such standards and the development of standards and nomenclatures.

The development of standards is a complex, lengthy process that depends among others on the availability of resources, especially experienced staff, to engage in such 'harmonization' efforts. Therefore, beyond in some advanced national statistical offices, harmonization of methods, classifications, technologies and guidelines is very much initiated and done by international statistical organizations. For these organizations comparability between regions and countries is a key condition, for example, for the monitoring of the progress of countries on the sustainable development goal indicators.

Work on standards goes beyond development and implementation. It also contains keeping them up to date. As described by Havinga [1] the majority of discussions in the United Nations Statistical Commission, in international working groups and task forces as well as in the international statistical organizations, are related to the development, implementation, and operationalization as well as to the evaluation and update/revision of the standards. Comparability due to diverging implementation of such standards is an often recurring theme.

Comparative statistics, used for factual comparison were (and still are) produced on a domain-specific base by global international organizations (for example UNCTAD see [2], FAO see [3]) and with a regional perspective by regional organizations like the UN regional organizations. The UN Economic Commission for Europe (UNECE) was – based on the rather advanced state of the European Countries' statistical systems – in the

eighties and nineties of last century even a forerunner in harmonization. Due to the legal base for these organizations, the standards developed were characterized by a normative non-mandatory character.

The growing importance of statistics in becoming a piece of basic information for international policy-making – initiating and monitoring policies – caused cross-national comparability to become so essential that the use of harmonized standards became obligatory and even legally prescribed. The European Union with a governance system very much based on the fair treatment of each member state, both on the financial contribution to the union as well as the use of common funds, has a reference to the use of statistics in its basic treaty. With the increase in the number of member countries especially since 2004, as well as the widening responsibilities and common goals, the role of statistics in the functioning of the European Union grew more important. As a result, the work in the European Statistical System on the development and implementation as well as making standards legally binding has been highly developed (Many of the official European statistics are part of the so-called '*acquis communitaire*').[7] In the EU, the culture of using statistical indicators has developed into a backbone of policymaking. International agreed standards have become the main 'language' and politicians and decision-makers almost blindly trust the comparability of indicators calculated according to these norms; and expect that these norms are properly applied.

The list[8] of international standards, guidelines, and recommendations (statistical classifications) has substantially grown over the years. Some of the standards exist for a longer period and are implemented in many countries. Such older standards also often have gone through several revisions. A well-known example is the System of National Accounts[9] (SNA). Revisions aim to update the standard to the most current technical and methodological level as well as to include all (new) relevant phenomena. Nevertheless, many countries, due to the complexity and resources needed, have difficulties implementing revisions. Consequently, the validity of comparisons can be affected by the use of different versions of a standard.

For more recent initiated standards, like the System of Economic and Environmental Accounts[10] (SEEA),

---

[6]For example the International Organisations with a Statistical department and/or an objective in their mandate on collecting and publishing statistical information, as these are organized in the Coordination Committee for Statistical Agencies (CCSA).

[7]https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:acquis.

[8]See: https://unstats.un.org/unsd/classifications/.

[9]See: https://unstats.un.org/unsd/nationalaccount/sna.asp.

[10]See: https://seea.un.org.

from the adoption some 5 years ago, only a few countries or regions can implement them fully. Depending on their complexity a certain number of years is needed to implement new standards. Thereby, implementation itself as well as the speed of implementation is very much dependent on resource issues.

Nevertheless, with the growing importance of cross-national comparisons in more and more domains, and the need felt by countries to provide and comply with the demand for high-quality statistics, the number of countries that want and factually adhere to the standards has substantially increased.

An important part of the work program of International statistical organizations is devoted to the outreach of these standards and methodologies including their introduction to less-developed statistical systems. Statistically more advanced countries are also called to support the implementation and use of the standards. Tools for support of the accessing and implementation of the standards are traditionally knowledge exchanges via training, or via programs and projects directed to implementing such standards and guidelines.

4. Biases in implementing the International standards and communicating on their use

In a period with a high prominence given to statistical indicators and statistics by the public at large and by policymakers, progress in countries and regions in a wide variety of domains is measured and communicated in terms of their score on indicators. Unavoidable are interpretations that rank, based on such scores, countries to do better or worse than others. Lists of countries (or regions) ranked on certain indicators are the most common representation of statistical information beyond national statistics.

But for these comparisons to be statistically valid (as opposed to culturally valid), the underlying standards must have been applied correctly and consistently. While for some indicators this is relatively straightforward, others are complex and open to interpretation. For example, the Gross Domestic Product of a country is based on the standards set by the System of National Accounts (SNA). Different forms of implementation and interpretation can lead to information that should be comparable, but in reality, can differ. Considering the importance of the indicators based on the SNA the discrepancies between the standard methods and those applied should be meticulously documented in metadata. However, this is not always the case.

Communicating about the quality of the data used, or meta-information on the methodology used is also dependent on the expertise, knowledge, and awareness of the statisticians involved. Statisticians might be hesitant to show openly that they cannot fulfill the requirements of a global standard. Or, there might be situations where policymakers have to be satisfied with some information rather than none. For several indicators, where the implementation of the standards is relatively straightforward, differences in implementation of the standard (be it by the methodology used, or the nomenclature applied) will be by definition rather limited. Though, many indicators are based on more complex calculation and production methods and require higher criteria concerning the quality and completeness of the underlying data.

There might be situations where, purposely the misfit between the standard and the available data, etc is not communicated. Statisticians might be hesitating to show openly that they cannot fulfill the requirements of a global standard. Or, there might be situations where policymakers have to be satisfied with certain information, some information might then be better than no information.

5. Biases in the contribution to the development of international standards

To achieve international agreement on any statistical standard involves compromise. What is economically, environmentally or culturally important in one country or region of the world might be of little relevance or importance in others. The trick is to devise systems that can be used for national purposes as well as for international comparisons without having to completely repurpose them. But for this to be achieved, all countries must have an equal opportunity to express their priorities and contribute to decision-making.

Most of the global statistical standards currently in use have been developed by working groups, task forces, 'city groups', and other expert groups. Resource issues (financial but also time) mean that it is often those countries with more advanced statistical systems that play the leading roles in such fora. The richness of the available data, expertise, and resources in these countries or institutions, as well as the demands made on them by their policymakers, may mean that the standards developed by such groups are biased towards their situations and requirements. Less well-resourced countries do not – or are not able to – have an input to

such discussions often until a late stage in the process, for example when the standards reach the UN Statistical Commission for discussion. This may be through choice, because of the lack of perceived relevance of the discussions to their situations, or because of lack of finance to attend meetings taking place around the world.

This was often happening with the assumption that the other countries, over time with the development of their statistical systems and societies, would join, based on development programs, the group of countries that can apply properly these standards and methodologies. An example of the influence of (financial) resources on the biased development of standards is that in international meetings exactly statisticians from poorer countries because of the lack of funding are not at the table when the statistical developments are discussed.

There may be an assumption that these countries would, over time with the development of their statistical systems and societies, adopt these standards and methodologies. But if the standards have not taken into account their particular needs and situations in the first place, this also seems doubtful.

An area that can be seen as an example of a bias based on hampering knowledge and experience is the measurement of SDGs. Experience shows that the methodologies for estimation of indicators developed by custodian agencies are not understood by all NSOs in the same way. Even if understood correctly, the statistical databases do not allow for meaningful estimation of indicators, etc.

6. Bias in the coverage of the international standards

The fact that most of the standards are developed based on expertise and knowledge of more developed statistical systems holds beyond the capacity issue another weakness. The voice of the not-so-developed statistical systems might not be heard and consequently, the specificity of these countries is not included in the considerations for the standard or methodology. The standards as developed are so by definition (concepts and operationalization) biased to a small group of developed statistical systems.

Applying a global statistical standard to produce statistical indicators, and referring to the standard that has been used in the methodology will lead the user of the cross-national comparisons to assume that they are valid representations of 'real' differences between the societies compared. Such a situation can be caused by the producers of the statistics, but can also be caused by the compiler of the cross-national comparison. So, there are several reasons why this assumption of a valid representation bears a risk to be not or only partially thru.

On the side of the producers of the statistics, the national statistical organization, the lack of awareness and/or expertise and knowledge of the standard, the lack of the needed data (sets), and/or the lack of resources can cause that the standard is not (wholly or accurately) applied in producing the indicator. When someone else than the national organization for statistics is producing the indicator, the same situation can appear, but also negligence (by a lack of knowledge or awareness) of the local situation can cause a misinterpretation of the applicability of the standard. But the compiler of the cross-national comparisons may also be unaware of the local situation which can lead to a misinterpretation of the data.

Also missing or unclear reporting on the extent that in the production of statistics for a certain country or region the standard is used according to its requirements might cause users of such statistics to be thoughtless to apply the results.

Being aware and having documented information about these biases and risks might lower the risk of the wrong use of this information. However, there is rather some (peer and political) pressure on the use of cross-national comparisons in international official statistics. This pressure might lead to cross-national comparisons that for the above-mentioned reasons are giving misleading (non-valid) outcomes.

7. Even when standards are implemented properly the cross-national comparisons might be meaningless

Even taking into account the above – the biased development of standards and the variable quality of their implementation – there is the risk that they do not properly represent the society in question. Certain variables and indicators might, from the national perspective, simply be not valid representations of society. Operationalizing a characteristic (of households, businesses, etc) according to the international proposed standard might not do justice to the local cultural and administrative differences. Due to this, policymaking might be directed onto other less relevant aspects.

Beyond the influence of lacking methodological expertise and available data, this biased development of

standards also risks that the voice of the less developed systems is not well taken on board. Specific variables and indicators might from the national perspective, simply be not valid representations of the societal characteristic. Operationalizing a characteristic (of households, businesses, etc) according to the international proposed standard might not do justice to the local cultural specificities, both in measuring as well in policymaking. Policymaking might be directed, due to the specificities, on other aspects, consequently, the cross-national comparison is of no use for national policy-making and-effectively also of relatively low use for international policymaking. There is a situation of different contexts.

All this brings us to the statement that standards and guidelines might look to be implemented and that the indicator represents a state of society, however, this is only on the surface, the situation below is very different. We might be comparing apples with peers! We risk that many cross-national comparisons in international official statistics are artifacts that by the unconscious use of the assumption that they are based on internationally accepted and agreed on standards can lead to erroneous and misleading conclusions.

8. What should we as statisticians do?

In conclusion, though the 'standards' are the backbone of modern official statistics it is valid to question if the frequent use of cross-national comparisons dismisses the cultural specificities of a country or region. It is also relevant to ask if the standards that are used to produce the indicators for cross-national comparisons are sufficiently implemented to allow for valid comparisons. And in general, one might question if there is a misfit between the emphasis on and practice of cross-national comparisons by international organiza-

tions and the attention to the level and awareness of the implementation of the standards used to produce the indicators on the country level.

From the above, it can be argued that cross-national comparisons based on global standards can be risky. Of course, as statisticians, we are fully aware of the statistical fallacies, invalid operationalizations, validity, and reliability issues. However, other circumstances in the development and the use of standards have to be taken into account when making cross-national comparisons. These circumstances are not always in our sphere of influence and vary from political arguments to general decision-making attitudes and procedures in use in a society where some countries/regions/organizations can have a stronger voice than others.

So – what should we, as statisticians, do?

References

[1]   Havinga I. Stylized facts of statistical standards of the Statistical Commission of the United Nations. Statistical Journal of the IAOS. 2022; 38(2). IOS Press, Amsterdam.

[2]   MacFeely S, Peltola A, Barnat N, Cristallo D, Chernova E, Hoffmeister O. The role of international organizations in statistical standards-setting and outreach: An overview of the UNCTAD contribution. Statistical Journal of the IAOS. 2022; 38(2). IOS Press, Amsterdam.

[3]   Bizier V, Gennari P, Kalamvrezos Navarro D. Role of international, regional, and country organizations in adapting to food and agriculture statistical standards and regional differences. Statistical Journal of the IAOS. 2022; 38(2). IOS Press, Amsterdam.

[4]   Bersales L. Standards in Social statistics: The capacity of national statistical systems to adapt to the international standards, problems, challenges'. Statistical Journal of the IAOS. 2022; 38(2). IOS Press, Amsterdam.

[5]   Kouadio H. Effectiveness of the outreach of official statistics standards and guidelines, methodologies and recommendations to developing statistical systems: A discussion. Statistical Journal of the IAOS. 2022; 38(2). IOS Press, Amsterdam.