



Economic and Social Council

Distr.: General
7 April 2022

Original: English

Economic Commission for Europe

Conference of European Statisticians

Seventieth plenary session

Geneva, 20–22 June 2022

Item 6 (a) of the provisional agenda

Reports, guidelines and recommendations prepared under the umbrella of the Conference: Measuring the value of official statistics

Measuring the Value of Official Statistics

Prepared by the Task Force to test and develop the framework for measuring the value of official statistics

Summary

This document is an abridged version of “Measuring the value of official statistics: testing and developing a measurement framework”, prepared by a dedicated Task Force consisting of the United Kingdom (chair), Armenia, Australia, Belgium, Canada, Czechia, Denmark, Finland, Hungary, Ireland, Israel, Lithuania, Mauritius, Mexico, Namibia, Netherlands, New Zealand, Poland, Romania, Slovenia, United States of America, Gulf Cooperation Council and an independent expert.

This abridged version of the document has been prepared for translation purposes. It includes the introduction, selected parts of chapter 2 on what is value and why do we want to measure it?, a section of chapter 3 with guiding principles for refining the measurement framework, and the full chapter with conclusions, recommendations, and future work, corresponding to chapter 6 in the full document. The numbering of the chapters and sections in this abridged version does not correspond to the numbering in the full document.

The full text of the document includes also chapters on a new proposal for understanding value; reviewing the measurement framework; and a results-oriented approach to value measurement called a Results Map; as well as a reference list, and annexes containing the original measurement framework, the complete evaluation of measures and a country case studies. The full document was sent to all CES members for electronic consultation in March 2022 and is available on the Conference webpage at <https://unece.org/statistics/events/CES2022>. A summary of feedback received during the consultation will be provided in document ECE/CES/2022/3/Add.1.

Subject to a positive outcome of the consultation, the CES plenary session will be invited to endorse the document.



Contents

	<i>Page</i>
I. Executive summary	3
A. Key messages	4
II. Introduction	6
A. Background.....	7
B. The measurement framework	7
C. Evaluating, testing and refining the framework.....	8
III. What is value and why do we want to measure it?	8
A. Introduction	8
B. The changing face of the value of statistics in the context of the Covid-19 pandemic.....	9
C. Delimiting concepts.....	11
IV. Reviewing the measurement framework	15
A. Guiding principles for refining the measurement framework.....	15
V. Conclusions, recommendations and further work	16
A. Conclusions	16
B. Recommendations	17
C. Further work.....	18

List of figures

Figure 1:A conceptual map of value within different schools of thought.....	13
---	----

I. Executive summary

1. Any attempt to quantify how valuable official statistics are must begin by asking what is meant by value, and whose perceived value matters.
2. This report reviews potential ways of measuring value, supported by case studies demonstrating their use and suitability (or non-suitability) for assessing the value of official statistics. The overarching conclusions of the document are:
 - Consumer-based approaches to value result in a different set of potential indicators than those arising from a more traditional, ‘production-based’ approach to value.
 - While production-based indicators can be very useful for operational and management purposes (including areas like quality and budgeting), they do not necessarily reflect the value of our outputs in the sense understood here.
3. Moving towards a consumer-centred approach to measuring value—and, by extension, to creating, maintaining and improving that value, which is the ultimate goal—necessitates a fundamental shift in direction.
4. This document offers a suggested method to navigate this change, which entails using a ‘Results Map’ to define a clear path to achieving the central goals of official statistics; working outwards from core strategic goals to measurable outcomes, to quantitative indicators of value.
5. Official statisticians are driven by the conviction that their products, underpinned by the [Fundamental Principles of Official Statistics](#), are uniquely valuable and essential for evidence-based decision-making, political accountability and democracy. But there is an increasing awareness within the community of official statistics producers of a disconnect between their own understanding of the benefits and value of their products, and the public perception of the value that official statistics can offer. It is not merely a matter of ‘proving our value’. Official statistics are not the only source of statistical information. Consumers justifiably will choose, use and trust official statistics over other sources only when their own weighing of the pros and cons leads them to do so. Governments and other funders want to know if the resources invested in official statistics offer a good return on their investment. Official statisticians want their commitment to the Fundamental Principles, and what they view as the value resulting from this, to be recognized and understood.
6. Driven by all these motivations, in 2015 a Task Force on Valuing Official Statistics was established by the Conference of European Statisticians (CES), resulting in the 2018 [Recommendations for Promoting, Measuring and Communicating the Value of Official Statistics](#). The Recommendations promulgated a measurement framework comprising a set of proposed indicators for measuring value, and some suggested methods for producing them. In concluding their work, the group suggested that ‘pathfinder’ National Statistical Offices (NSOs) should pilot test the proposed measurement framework, with the support of the United Nations Economic Commission for Europe (UNECE), and should share their findings and experiences with the wider international statistical community.
7. The present work originated from plans to undertake such pilot testing, via a Task Force established in 2019 for this purpose. Yet two things became clear very early on in the exercise:
 - First, the proposed framework inherited from the 2018 Recommendations was a brainstorming exercise, not yet sufficiently developed to be able to ‘pilot test’ without first refining and developing it;
 - Second, the whole exercise of developing and testing a measurement framework rested on the untested assumptions that it was already clear what ‘the value of official statistics’ meant, and for what reasons the official statistics community would wish to produce such measures of value. In fact, as the present document shows, these two assumptions needed to be examined and deconstructed in depth before any ‘pilot

testing' could be meaningful and before a useable measurement framework can be developed.

8. This publication presents the findings of the Task Force. The main argument put forward by this group is that 'value' is determined by the customer. Therefore, any attempt to quantify how valuable official statistics are must begin with an investigation into what it is that people value. A measurement framework comprising indicators of how well official statistics live up to the standards set by the statistical community themselves, no matter how laudable and universal those standards and no matter how detailed and carefully crafted the indicators, is not, in fact, a framework to measure the value of official statistics.

9. The report documents the process followed by the Task Force to reach this core argument, including a detailed review of the measurement framework, which found that a significant share of the original measures cannot be proposed as indicators of customer-perceived value. Many are still useful for operational reasons, including quality assessment and budgeting, but a much-reduced number are retained as having potential as indicators of value. Amongst those retained are a number of longstanding measures that show or have the potential to show value, and there is ample evidence of methodologies becoming more and more developed in the quest to promote and improve official statistics. Many countries have undertaken or are undertaking targeted work to evaluate and assess their impact, covering a range of objective, subjective and monetary angles.

A. Key messages

10. The key messages arising from the two years of deliberations and investigations by the Task Force are the following.

1. Before we can measure value we need to know what it is

11. Just as with any topic in statistics, an essential precursor to measurement is defining concepts and delineating what is to be measured. 'Value' is a fuzzy concept with a range of meanings, making it hard to give a precise definition - but this only makes it all the more necessary to be explicit about what we are talking about.

2. We measure value with a view both to proving and to improving

12. Developing a meaningful framework for measuring the value of official statistics relies on us knowing why we want to measure value at all. One reason is to prove to others that what we do has worth and is a good use of public resources. Another reason is to enable us to monitor the effectiveness of our efforts to increase our value. Both reasons are valid.

3. The only way to determine what people value is to seek their views

13. We should not assume that we know what people value; or that what we think is important is the same as what they think is important; or that if we explain well enough why we think things are important, people will come to share our views. To really know what elements comprise the value of official statistics, we need to ask. This doesn't have to mean direct questioning, which could be poorly understood, but it does have to entail some method of probing users to identify their criteria.

4. Value ≠ values

14. Value, like beauty, is in the eye of the beholder. What makes official statistics valuable is not for us to say (unless we ask—see previous point). We can state our values, the things that drive and motivate us and the reasons that we do what we do. We can state the properties of official statistics that result from the values we hold. And we can even state why we think users or society at large 'should' find these properties valuable. But there is no automatic link between values and value. In the end, users will value what fits their needs and it is they who define those needs, not us. Relatedly, communicating our values is no guarantee that this will convince anyone of our worth. Like the foreign tourist who speaks louder and louder in their own language in the hope of being understood, we risk alienating our users by simply stating

our values and hoping that people will ‘get it’. More helpful would be communicating how we fulfil *their* criteria of value.

5. Value ≠ quality

15. In official statistics, quality is already a well-defined concept with clear and agreed dimensions. It amounts essentially to 'how good our statistics are'. Although one of the dimensions in standard quality frameworks is relevance, which is defined as meeting users' needs, the majority of other quality dimensions are aspects of the statistics themselves which can be relatively objectively measured within an NSO. In contrast, value cannot be determined without reference to the perception of the one doing the valuing. It is inherently subjective. For some users, it may be synonymous with quality, if quality is the main thing they value, but for others it will encompass more intangible aspects including relationships and opinions. For example, for some users, the simple fact of coming from an NSO and the gravitas that offers may be enough to make official statistics more valuable to them than unofficial ones, even if all other aspects of the statistics could be held equal.

6. Measures of value are not (necessarily) the same as measures of quality nor of adherence to values

16. Since value is not a synonym either of values or of quality, a framework for measuring value will not consist of identical indicators to those found in quality frameworks, nor will it consist of measures of how well we uphold values and embody the Fundamental Principles. But there may be some overlap, since for some users these things are indeed central components of their perception of value.

7. Testing and refining the measurement framework has highlighted underlying issues with the exercise

17. The mandate of the Task Force was to test and refine the framework as formulated in the 2018 Recommendations. The content of this publication demonstrates that this was done. But a principal finding is that the exercise of formulating a measurement framework "from the outside in", starting with available indicators and testing them to see whether they provide a good picture of the thing being measured, follows a flawed logic. The resulting refined framework proposed here, therefore, both sets aside a significant number of measures from the framework because they were found not to provide a good picture of the thing being measured, and contains notable gaps, for which new measures should be developed as this work is taken forward.

8. Excluding measures from the framework does not mean they are unimportant

18. The work undertaken by this Task Force involved reviewing each of the proposed measures included in the 2018 framework, in light of the collated experience of 18 countries and organizations who shared their views. This review resulted in a significant number of those measures being excluded from the framework, for a range of reasons: because they are not considered to be actually indicative of value; because they are not quantitative indicators; because they are not well-grounded in a logical pathway from intended impact, to outcomes, to measures indicative of achieving those outcomes; because the behaviours implied by tracking that measure are not desirable; because it is not evident what a 'good' level of the indicator would be or it does not suggest a target towards which to strive; and/or because there is no obvious monotonic relationship between the indicator and the 'value' it purports to measure.

19. Nevertheless, excluding a measure from the framework for any or all of these reasons does not mean there is no point in producing it. It just means that it is not proposed by this group as a potential indicator of the value of official statistics. It could be an indicator of something other than value. There may be many other reasons to produce the measure. The case studies illustrate a number of examples of this.

9. A framework for measuring value can include things we don't currently know how to measure

20. There are many possible reasons to exclude a suggested measure from the measurement framework, but the difficulty or current impossibility of measuring it should not be one of them. If we think something would be a good indicator of value but we don't currently have a means of quantifying it, that should not be a reason to throw it out. Indeed, we of all people, as statisticians, should not fall into the trap of saying that if something is hard to measure then we won't even try: instead we should attempt to develop a way to measure it! This being said, of course, the converse is also true: just because a quantitative measure of something *does* exist does not mean it necessarily is relevant and should be used.

10. A useful approach to developing a value-measurement framework is to use a Results Map

21. In addition to reviewing and assessing the 2018 measurement framework, the Task Force considered an alternative means of developing a measurement framework. Instead of beginning with possible indicators and testing them to see what works, this approach works in the opposite direction. It begins by defining the central goals of official statistics, then works outwards through organizational strategies that help achieve these goals; to tactical outcomes; to measurable areas for which indicators can be developed; and finally to the indicators themselves. While much more time-consuming and with the real possibility that not all proposed indicators can (yet) be easily measured, the approach has the major strength that every resulting indicator in the framework traces a clear path from the central goal.

11. The refinements proposed to the measurement framework are only a first step: this Task Force proposes a fundamental shift in approach to take the work forward

22. The reviewed and refined framework proposed in this publication is based on assessing the measures suggested in 2018. Following through the Results Mapping process in depth, either on a country-by-country basis or collectively on an international level to develop a generic measurement framework capable of being adapted to different countries' circumstances, is a task beyond the mandate of the present Task Force; but one which this group argues could be extremely useful to help inform ongoing efforts to prove and improve the value of official statistics. Rather than pilot testing individual indicators of value, therefore, this group proposes that future work centre on pilot testing the Results Map procedure, drawing heavily on user (and non-user) consultation, to guide official statistics in its endeavours to offer value to society.

II. Introduction

23. "What makes official statistics valuable to you?" More than a thousand people viewed this question when posted by a statistics professional in a highly engaged network on LinkedIn. Yet only two offered substantive answers.

24. Posted on official social media accounts, the question prompted support in the form of 'likes' and 'retweets', but not one reply.

25. What does this tell us? That those who interact with official statistics via social media platforms cannot think of a single thing that makes our work valuable? Should we just pack up and go home?

26. The work presented here argues that official statistics, as an industry, is most likely immensely valuable to users and to society as a whole, but that our introspective habits built up and institutionalized over many years have led us to take this value for granted, leaving us without a clear means either of proving or of improving it.

27. One reason why people do not, or perhaps even cannot, answer the question posed above is that it is too abstract. No industry, no provider of any product or service, would expect their customers to be able to articulate directly and consciously the range of criteria they apply, or the relative weights they assign to these criteria; for example when choosing a car, a pair of jeans or a brand of cookies. If it were this simple, market research in any field

would be a much smaller operation. In fact, it is a whole industry based on subtle and complex ways of eliciting information from people about the criteria they apply, often subconsciously, and how they relate these to value, whether in terms of value for money, time, prestige, or other factors.

28. So it must be for official statistics.

29. First, ascertaining what criteria our users (or potential users, passive users and non-users) apply when determining the value they assign to our products and services is a complex operation, because such determination is rarely conscious or explicit, and because it is almost certainly not fixed but varies across user, use, time and space.

30. Second, the task of ascertaining these criteria is not one that can be done purely internally without consulting those very users—even though such consultation is challenging, and as the social media example above illustrates, cannot be achieved by simple direct questioning alone. If we opt instead to come up with our own components of value based on what we assume to be valued by users, we risk identifying the wrong criteria and then being led by these towards erroneous goals.

31. Third, even once the criteria—the constituent elements that together result in a user’s perception of value—have been established, operationalizing them in a way that permits us to monitor, compare, set targets and assess progress towards them, is an equally large and multifaceted task.

A. Background

32. This section has been omitted from the abridged document.

B. The measurement framework

33. The 2018 Recommendations put forward a framework made up of three components, or kinds of measures of value:

(a) Observable or objective indicators (reflecting actual use of statistical products; measures of quality assumed to be adding value; and/or demonstrating adherence to the Fundamental Principles of Official Statistics);

(b) Subjective indicators (covering perception, trust, support, satisfaction etc.); and

(c) Monetary valuations (quantifying the impact of statistics in monetary terms and/or weighing the value of outputs against the cost of inputs).

34. In general terms, the objective component consists of things which could be measured with relative ease from existing sources, such as download counts, numbers of citations in various different types of media, social media interactions, etc. The inclusion of these as proposed indicators of value rests on the assumption that when people view, download or use our statistics or mention our work, this is an indication that they find our work valuable.

35. The subjective component comprises measures which would primarily be obtained from user satisfaction surveys—either dedicated full annual surveys, or continuous mini surveys on web pages. This component offers crucial insights into user confidence and trust in official statistics, and into how useful, relevant and accessible they are found to be by users.

36. The monetary component merits a heading of its own due to its communicative power, even though it is not in the same conceptual or logical category as subjective and objective measures (it is not a mutually-exclusive three-way classification). Much of the discussion around the value of official statistics relates to the desire to investigate whether official statistics are good value for money, as a public good resourced by public money. Monetary measures of value are the most complex to produce but particularly easy to communicate. Given the unique position of NSOs being in the field of quantification, there is an expectation that they should be able to defend in quantitative terms the value that they add.

37. Within these three components, the 2018 framework suggested wide range of indicators or measures, grouped into sub-components (e.g. the objective component is divided into indicators of quality, transparency, use and relevance; the subjective component includes, among others, the sub-components ‘awareness of brand and message’ and ‘user support’).

C. Evaluating, testing and refining the framework

38. The CES work on measuring value was established with the intention of ‘pilot testing’ the measurement framework. But early on in the activities of the group two things became clear. First, it was found that that the framework was not yet something that could simply be pilot tested. The loose constellation of measurement ideas proposed in the 2018 Recommendations was an early brainstorming exercise, not yet sufficiently sophisticated to be operationalized into actual measurable indicators. ‘Number of tailored services by user group’, for example, or ‘share of users whose information needs were met’ are far from being sufficiently clear to be able to produce figures without first establishing a variety of necessary definitions and a standard methodology.

39. Second, the group found that some of the items in the proposed framework were not, in fact, measures of value, while others that perhaps ought to be included in a comprehensive consideration were absent from the framework. Indeed, the whole exercise of developing and testing a measurement framework rested on the untested assumptions that it was already clear what ‘the value of official statistics’ meant, and for what reasons the official statistics community would wish to produce such measures of value. In fact, as this review shows, these two assumptions needed to be examined and deconstructed in depth before any ‘pilot testing’ could be meaningful and before a useable measurement framework could be developed.

40. The group therefore expanded its aims, to include both testing and developing the measurement framework. To develop it, they first needed to achieve clarity and consensus on what value actually means, and why NSOs would or should want to measure it. On this basis, they designed a more coherent and complete approach to measuring the value of official statistics. In combination with detailed input from each participating country about their own experiences in attempting to produce measures of value in the various aspects of this framework, and complemented by a rich set of case studies, the Task Force was then able to put forth generalized recommendations for NSOs on whether, why and how to apply the framework.

41. This remainder of this chapter has been omitted from the abridged document.

III. What is value and why do we want to measure it?

A. Introduction

42. In today’s data environment, it is no longer enough to take the value of official statistics for granted: to assume such value exists and that stakeholders share our perceptions of it. Big Data and the accompanying boom in statistical products has accelerated an evolution in customers and in the products that they demand, as policymakers, businesses and the general public increasingly look for timely and reliable data to guide their decisions. This growth is combined with an increased ease of access, greater use (and misuse) of statistics, and circulation via social media platforms that makes the data market surrounding NSOs more competitive than ever before.

43. The Covid-19 pandemic has deepened and accelerated these trends, acting as a catalyst for changes that were already underway, not only in the production of statistics but also in the way that NSOs assess and reflect upon the value of their contributions to society. Section 0 below shines a spotlight on the ways in which the pandemic has affected society’s and NSOs’ understanding of value.

44. Given these changes, NSOs have an opportunity to measure the value of official statistics and to communicate it in a compelling way to stakeholders including to the wider public. Currently, NSOs are broadly recognized as independent bodies, whose products and services can be trusted by society. But they should not be lulled into complacency by relying on institutional trust or prestige to navigate the new statistical landscape. Nor should they underplay the value of official statistics, particularly as they look to defend their budgets or advocate for more resources.

45. The ability of NSOs to review the value of their outputs from the perspective of users, and to set improved value targets that respond to user needs, will define the continuation into the future of the traditional high esteem and relevance they currently enjoy.

46. In addition to highlighting the positive impact of the work of NSOs and the larger National Statistical Systems (NSSs) on society, the process of measuring the value of official statistics will intensify dialogues with users, potential users, and ‘passive users’ (anyone who benefits in some way from the use of the official statistics, from their use in decision making; they may not consider themselves users of statistics but they benefit from their existence). This in turn will enable NSOs to monitor the effectiveness of their efforts and support product development to better meet the needs and interests of users.

47. Private statistical providers track the value of their products by contrasting input efforts against revenue, as their survival is tied directly to the responsiveness of their customers. Prices reflect willingness to pay and patterns of consumption, which are aligned with trends and deeper changes. NSOs differ both in their core function – to provide their countries with data that informs decision-making in every layer of society – and circumstances – being publicly funded and forming part of the national institutional fabric. Therefore, as with any public good, the process of weighing the value of their contribution must expand beyond a purely monetary approach. Ultimately, the value derived by users depends on whether their needs are met in order to make well-informed decisions with reasonable confidence.

B. The changing face of the value of statistics in the context of the Covid-19 pandemic

48. Within days of the World Health Organization (WHO) declaring a global health emergency, a change began to emerge in the way that statistics and data were being used and perceived across the world. Daily press briefings from the WHO and other international organizations, national governments and news bulletins began to centre around numbers and graphs; people who had not needed to look at or interpret a graph in years found themselves studying trendlines and axes; whole editorial news articles were devoted to picking apart the nuances of different indicators of mortality, hospitalization and positive cases.

49. As the main custodians of official data on these questions, this threw NSOs into uncharted waters. They were expected to provide and stand behind figures, more rapidly than ever and against a level of scrutiny never before witnessed. Official statistics producers have always been proud of the fact that they are transparent, scientifically robust and politically independent. Some users, in the past, may perhaps not have placed especially high value on these characteristics, but as national responses to the pandemic often became politicized, these qualities suddenly took on a central importance for many. At the same time, the relative importance of various quality dimensions—timeliness, punctuality, accuracy, accessibility and others—in the eyes of users underwent shifts, as the nature of the users and uses transformed.

50. This is not to say that the meaning of the term value has changed. The idea of value being shaped by “fitness for use” remains unaltered. But the uses to which statistics are now being put and the demands being made of the offices producing them are in many ways very different from what they were before the onset of the pandemic. Policymakers, scientists and the public have needed information urgently to inform their rapid-response decisions related to the progress or impacts of the pandemic; and as such, they may find themselves placing greater value on the timeliness of statistics to fulfil these urgent needs, potentially placing

less emphasis on precision and level of detail, often touted as the ‘gold standard’ for official statistics.

51. Many NSOs were already starting to assess the value of their outputs prior to the Covid-19 pandemic, responding to the proliferation of statistical products that have become available in the marketplace. Today, however, we are witnessing how the perception of value changes in line with circumstances and the needs of the user or customer, making it clearer than ever that these value assessments need to be regular and ongoing, not one-time-only evaluations.

52. For example, with the onset of the pandemic, governments were confronted with very different data needs to support their pandemic response planning. Gross Domestic Product (GDP) metrics are of course still required to measure the economic impact, but a highly precise, quarterly product with a six-month time lag is inadequate in circumstances such as lockdowns where changes are large, rapid and subject to major short-term fluctuations. Instead, ‘flash GDP’-type products have evolved. These products feature much higher frequency and timeliness than traditional measures. Whilst these products diverge somewhat from the traditional value model of official statistics, they respond directly to customer needs (and their perception of value) for dynamic decision-making.

53. A further example of the way that the constituent aspects of the value of statistics has changed, is the way in which demand shifted towards compact, ‘pre-digested’ statistics. Products that were easy to interpret, based on dashboards, visualizations, storytelling and simple graphics with clear tag-lines gained prominence. This corresponded both to increasing diversity of users (with more non-expert and less statistically literate users among them), and a need for more rapid processing even among more expert users such as journalists and civil servants who are used to working with numbers but lack the time to process them themselves. Many NSOs introduced their own national Covid dashboards or portals¹, and Eurostat developed the [European Recovery Dashboard](#), which was launched in December 2020.

54. In almost direct contrast to this heightened value of aiding user interpretation through increased pre-processing, a parallel shift took place. For some users, the availability of detailed, transparent, verifiable metadata was what gave official statistics the most value in the early days of the pandemic. For many, the special value of official statistics as compared to other sources derived from the ability to check sources, understand concepts and definitions, investigate the reasons for differences in mortality as measured in different ways and the reasons for certain peaks or troughs in infection rates, unemployment rates, testing frequency etc.

55. A final major shift in the value of official statistics—understood as an industry, system or NSO rather than in terms of individual statistical products or services—brought about by the pandemic is the greatly increased importance of official statistics’ role in safeguarding correct and fair use of statistical information. The role of chief statisticians, chief economists and statistical regulatory bodies in calling out deliberate or accidental misuse has always been a unique selling point of official statistics. The rise of Covid-related misinformation, combined with widespread mistrust of governments and politicization of Covid response policies, led NSOs rights into the spotlight as ‘custodians of statistical facts’. Extra efforts have had to be employed to ensure public understanding of the political independence of NSOs. From the point of view of a member of the public looking on as politicians debate their points using statistics, the knowledge that the figures being cited are produced independently from those citing them, and the knowledge that egregious misuse of statistics will be publicly called out, is likely to be more valuable than ever.

56. These shifts seen in users’ valuation of the different aspects of official statistics can, and must, lead to concomitant shifts in the focus of NSOs to ensure their work continues to be valuable. The specific changes identified might well be short term. Some of the new users may indeed gradually move back to being non-users or passive users, while the regular and expert users may eventually return to a situation where they are more interested in the degree of detail and accuracy than in how rapidly they can obtain preliminary figures. But what will remain as a permanent impact on official statistics is the realization that *what makes statistics*

¹ For example, [Poland](#), [Ireland](#) and [the Netherlands](#), among many others.

valuable is neither immutable nor universal. Statisticians are eager to measure the ‘quality’ of their products but (as discussed in section 0), quality as understood by statisticians is not a synonym for value.

C. Delimiting concepts

57. Statisticians know well that conceptual clarity is key to producing meaningful, comparable and interpretable measurements. This is especially true for terms originating in everyday language and now used in a specialized context, as in the case of the word ‘value’. Hence rather than assume a common understanding of the concepts that will serve as building blocks, the following sections aim to establish the terminology of this report.

58. The discussion differentiates what value is from what it is not; reflects on the characteristics that differentiate official statistics from other statistics; and explores the actors that make today’s data market so sharply competitive. The chapter aims to clear a path towards the factors that determine the elements of value, which NSOs can measure and set as targets for themselves.

1. Value

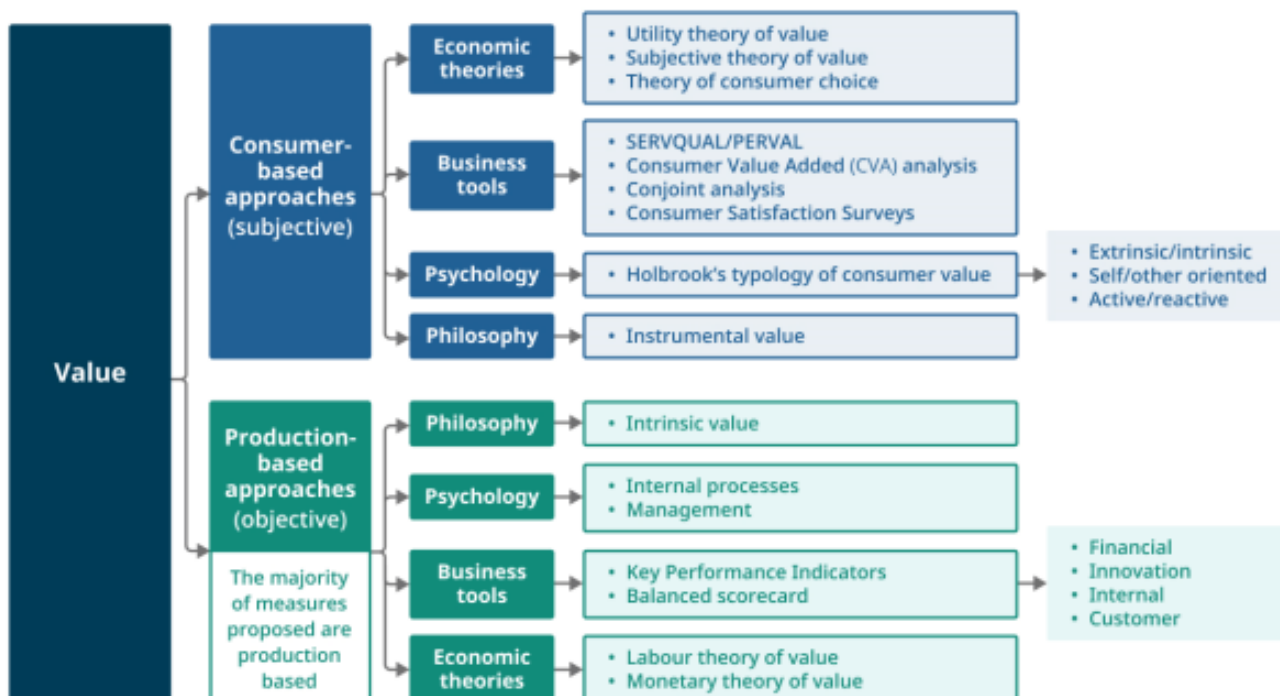
59. The word value can be understood in many ways and has been unpacked and defined in a variety of schools of thought, as illustrated in Figure 1, a concept diagram developed by the Task Force to map out the principal strands of thinking. While no attempt is made here to detail all of these approaches in an exhaustive manner, it is important to note that (as the diagram illustrates) there is a major bifurcation between production-based and consumer-based understandings of value. A production-based understanding of value is associated with the scarcity or abundance of a product (a good or service). This gives rise to the exchange or market value, i.e. the price, of that product. When used in this way, the word ‘value’ refers to something objective, determined by the market. A consumer-based understanding of value, in contrast, is associated with cultural traditions, human emotions and memories, and is therefore subjective and variable. Value understood in this way is not an intrinsic quality of any one thing: it is subjective, always perceived by someone within a context. Since official statistics are not a simple market good, their value cannot be understood by an objective consideration of price, but rather must be examined through the second, consumer-based lens. As the preceding Task Force on the Value of Official Statistics stated, it is rare for a product or service to be timelessly valuable. Therefore, value is a dynamic quality that has to be assessed and re-built continuously (UNECE 2018).

60. Terminologically speaking, value is what is known as a fuzzy concept. As such, it is first necessary to specify a determined context for what is meant by value, and then identify its component elements before measuring them. The present context is an assessment of the value of official statistics, from the perspective of users, for the demonstration and enhancement of that value by NSOs. The elements of value are determined by the qualities that users seek in statistics to inform their decision-making.

Figure 1
A conceptual map of value within different schools of thought

A conceptual map of value within different schools of thought

An objective of this report is to encourage a shift from how statistical organisations think about the value of official statistics, from production based to consumer-based approaches as value is determined by the customer.



2. Value and values

61. A major trend in current discussion among chief statisticians and other key players in official statistics at present is to refocus attention on, and reflect on the meaning of, the ‘core values of official statistics’. To a large extent these core values centre on the UN [Fundamental Principles of Official Statistics](#), as well as the ways in which these are enacted by NSOs in their work.

62. There is a very important distinction to be made between the value perceived by users and these values that NSOs appraise in themselves as institutions and in their own products and services. They are intermingled concepts that have important points of interaction, but they are not fully interchangeable.

63. First, while the value of official statistics changes as decision-making users react to the world and as social conditions evolve, the core values are set in place, designed to remain stable above temporary circumstances (as are the Fundamental Principles: hence the use of the word ‘fundamental’ to emphasize this continuity).

64. Second, value, like beauty, is in the eye of the beholder. ‘Values’, on the other hand, are things that we hold dear to ourselves, that shape our attitudes, motivate us, determine our priorities and guide our behaviours. For official statistics, these core values shape the way we see ourselves and the way we would like others to see us.

65. A person’s own deeply-held values may or may not align with what others find valuable in that same person, however. For example, someone might consider honesty and integrity to be the things that drive them, but ask others ‘what do you value in this person?’ and they might say their efficiency or good humour. There is no contradiction, it is simply the case that the two words refer to different things. That said, the values might lead directly to, or play a role in creating, the value: the values of honesty and integrity may lead the person to work hard, resulting in the efficiency that others value. The same is true in official statistics: our core values underlie and motivate what we do, and, we hope, result in features

of our work that society values. But we cannot assume that they are, or should be, the main or only things that society wants from us.

66. NSOs would dearly love for their core values to be known and understood by society at large. Efforts to better communicate the core values and the Fundamental Principles are ongoing both in the international statistical community and in individual countries. Underlying these efforts is an assumption that if they are known, then official statistics will be better appreciated, or valued.

67. But this assumption makes a logical leap that may not be entirely founded. Just because someone understands your values and what drives you, this does not mean that they will necessarily share your view. It might influence their perception of you, but it might not. Indeed, it might even have a negative influence. When we learn a person's motivation for their behaviour in daily life, this might lead us to hold this person in higher regard, or quite the opposite, depending on our own point of view. By the same token, simply ensuring that the core values and Fundamental Principles driving official statisticians are widely known, laudable as this goal is, is not enough. The present Task Force argues that communicating the values and how they add value is essential, as we cannot assume that this latter part will simply be deduced by society. And to do this, we must find out—from members of society—what constitutes the value of official statistics and hence how our work adds to it. This latter step is not something we can do alone. Introspection will get us so far, allowing us to examine our own beliefs and the central tenets upon which we have built our work, but without also examining the views of those assigning value to our work these efforts may be in vain (for more on this, see section **Error! Reference source not found.**).

3. Value and quality

68. For the purposes of this work, value must also be distinguished from quality. The official statistics community prides itself on quality. Indeed, quality is often touted as the hallmark of official statistics, the thing that sets our industry apart from statistics and data obtained from other sources. What do we mean when we say that official statistics have, or strive to have, high quality?

69. The International Organization for Standardization's (ISO) definition of quality is "the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs" (ISO 9000:2015). The United Nations Statistics Division (UNSD) (2019 p.7) puts this definition into more everyday language: "**Quality**: the degree to which a set of inherent characteristics of an object fulfils requirements. A simple definition is "fit for use" or "fit for purpose". It is the users' needs that define the quality. Different users may have different needs that must be balanced against each other."

70. The [United Nations Fundamental Principles of Official Statistics](#), developed in the UNECE region in 1992 and subsequently adopted as a global standard by the General Assembly, describe the core precepts that underpin all the work of national statistical systems: not only statistical products and services themselves but also the processes, perceptions, relationships with government and with stakeholders, and ethical standards.

71. Quality Assurance Frameworks (QAFs) and related guides on their application, such as the [United Nations National Quality Assurance Frameworks Manual for Official Statistics](#) and the [Quality Assurance Framework of the European Statistical System](#) and various national and international adaptations of these², are very closely linked to the UN Fundamental Principles. They operationalize quality as defined above in terms of a set of criteria or dimensions, grouped into those which refer to the institutional environment, those related to the statistical production process, and those which are concerned with the statistical outputs themselves. While the specific dimensions included in these frameworks can differ slightly across different countries, organizations, data types and product types, in general they comprise the following (UNSD 2019 pp.7–8):

- (a) Relevance: the extent to which the statistics satisfy the needs of the users;

² e.g. Statistics Canada, Statistics Sweden, IMF, Organisation for Economic Co-operation and Development (OECD).

(b) Accuracy: the closeness of estimates to the exact or true values that the statistics were intended to measure;

(c) Reliability: the closeness of the initially estimated value(s) to the subsequent estimated value(s) if preliminary figures are disseminated;

(d) Timeliness: the length of time between the end of a reference period (or date) and the dissemination of the statistics;

(e) Punctuality: the time lag between the release date and the target date by which the data or statistics should have been delivered;

(f) Accessibility: the ease and conditions with which statistical information can be obtained;

(g) Clarity: the availability of appropriate documentation relating to the statistics and the additional assistance that producers make available to users;

(h) Coherence: the ability to reliably combine statistics and data sets in different ways and for various uses. Consistency is often used as a synonym for coherence;

(i) Comparability: the extent to which differences in statistics from different geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the statistics.

72. It is possible, then, to use these standardized QAFs to assess the quality of anything an NSO does, ranging from an individual statistical product up to a statistical system as a whole.

73. So why not just rely on these frameworks? Why do we also need a framework to measure value? The simple answer is that quality and value are not synonyms.

(a) *The difference between quality and value*

74. The two terms, and the concepts they imply, are of course very closely interlinked and significantly overlapping. Parts of the two concepts, and their component elements, are at times even identical. Adding to the complexity of untangling the two concepts is that the extent to which they overlap differs according to the specific user and use.

75. The principal difference between quality and value is the perspective from which they are construed. Statistical quality is, at its core, a concept designed for the statistical office in order to monitor and assure the standards of the work they conduct and the products they disseminate. As such, even though the definition of quality given above is ‘fitness for purpose’, a large majority of its dimensions can be understood and even measured without reference to the user and without recourse to investigations of users’ perspectives. Of the nine quality dimensions in the UNSD framework, only one, relevance, absolutely necessitates explicit reference to the user and their intended use (a second, accessibility, can be interpreted in terms that depend on characteristics of the user such as their degree of statistical literacy or on physical and mental capacities). This means that with the exception of the relevance dimension, quality assessments can be prepared on the basis of largely internal exercises.

76. In contrast, value, as understood by this Task Force, is intended to be entirely customer-defined: very similar, in fact, to the relevance dimension of quality, but also comprising other, broader aspects (see **Error! Reference source not found.** for a discussion of these aspects). A framework for measuring that value may comprise elements that are quantifiable internally (everything from download numbers to customer service response times to timeliness of releases), but only a user can say whether these elements are actually components of value for them. Quality, in essence, is the ‘degree of excellence’, while value is the subjective assessment of that quality that makes something desirable.

77. An example can help to illustrate this difference. A child has a favourite soft toy. It is a rather poorly-made, cross-eyed, second-hand teddy bear. Whatever realistic criteria one might use to define the ‘quality’ of a soft toy, this toy does not meet many of them. But it is certainly of high value to the child. She loves it. The softest, handsomest, most well-stitched teddy bear with a silk bow around its neck would never meet her criteria for a valuable teddy bear. So it is with quality and value. For some children, the criteria will align entirely: some

will love the ‘perfect’ silk-bowed bear. For others, the criteria will intersect, with certain being features of both quality and value (softness, for example, might make bears both higher quality and more valuable for many children). But for others—such as this child—the criteria she uses to decide that her teddy is valuable have next to nothing in common with those we might use to say that it is high quality. And we have no right or power to suggest that her criteria are ‘wrong’, that she should change her mind, nor should we have any interest in doing so. We can communicate quality, but we cannot make users value it.

(b) *So which matters?*

78. Quality frameworks are designed to ensure that the statistics produced by NSOs are good, or ‘high quality’. They have a clear and obvious purpose and the work of this Task Force does not in any way intend to detract from the concept of quality, or from well-established procedures to define, assess and improve it. But quality is not a substitute for value. Both need to be measured, for different reasons. We measure quality to ensure that we are producing the best statistics we can, and we measure value to ensure we are doing what people want. A restaurant producing excellent haute cuisine will not do well in a town where most people want fast food. The successful restaurateur conducts market research to determine the wishes of their potential customers, and so must NSOs.

79. This remainder of this chapter has been omitted from the abridged document.

IV. Reviewing the measurement framework

A. Guiding principles for refining the measurement framework

80. In reaching their decisions on measures to be retained in the framework, the Task Force was guided by the following principles:

(a) Measures should be clearly indicative of the value of official statistics. As discussed in section 2.3.3, this is not simply a synonym of quality. Therefore, even though a measure may very obviously be a good measure of statistical quality or of adherence to core precepts of official statistics, this does not necessarily mean that it merits inclusion in a framework for measuring value;

(b) Measures should be (at least theoretically) quantitative and have a monotonic relationship with the aspect of value being measured. That is, a greater measurement indicates more value and a lower measurement indicates less value, even if the relationship is not necessarily linear;

(c) Measures should lend themselves to the formulation of actionable targets, i.e. it should be evident what a ‘good’ level of the indicator would look like, and it should be possible to envisage how an NSO could take action to harness the information gleaned by producing the measure, to inform some action or behaviour that could alter the level of the measure (that is, that could improve the aspect of value being measured). Measuring something which is entirely externally caused may be of interest for various reasons (and may have important uses for the business operations of the NSO) but it would not permit any action to improve the value of official statistics based on the measurement;

(d) The potential unintended consequences of producing a measure should be considered. The act of measuring anything is never neutral. By measuring something we imply that we find it important, and whether deliberately or not we may turn it into a goal or target, which can affect the behaviour of those whose actions contribute to the thing being measured³.

³ It is essential to recognize that, even if an NSO has not explicitly framed a measure as a target, it may become one. Even if we say that it is ‘just a metric’, the act of measuring influences behaviour, ultimately impacting on the thing being measured. If a team is judged on a key performance indicator, they will strive to improve that indicator, perhaps even at the expense of other tasks. If an NSO

81. Measures identified by many countries as long-standing key performance indicators are evidently considered important and are therefore typically maintained in the framework.

82. As well as categorizing the measures in the framework into those that are retained and proposed measures of value versus those that are not recommended, the Task Force also attempted to distinguish between measures that are grounded in a production-based conceptualization of value and those grounded in a consumer-based conceptualization of value. While this distinction is by no means clear-cut (some fall into both categories, and for some the categorization is unclear), this undertaking does go some way towards framing the selection of measures of value in a way that could help NSOs to reflect on the basis for their selection of measures.

83. This remainder of this chapter has been omitted from the abridged document.

V. Conclusions, recommendations and further work

A. Conclusions

1. We cannot take value for granted

84. The point of departure of this work, indeed, the opening sentence of the terms of reference of the Task Force was “Official statisticians know that their products, underpinned by the Fundamental Principles of official statistics, are uniquely valuable...” The deliberations of this group over two years revealed that starting from a more self-critical perspective would be helpful—that is, starting out by asking ‘are we adding value?’ (and endeavouring to determine the nature and extent of that value), rather than assuming from the outset that we definitely are.

2. We must not conflate value with quality nor with adherence to our values

85. Both statistical quality and adherence to central values of the official statistics community are fundamentally important and help to define and drive us and shape what we do. But they are not synonymous with the value of our work to society, and we must not treat them as such. Measuring value entails measuring something other than the degree of conformity with existing quality dimensions or compliance with agreed community values.

3. Measuring value helps us both to prove and to improve that value

86. In reviewing the measurement framework it became clear that some countries have been producing certain measures routinely for some time—sometimes for many years. Several countries reported during the review exercise that the task itself helped them to reflect on why they were producing these measures and what purpose they were serving. In some cases, they were being produced simply because they were easy to produce. But the introspection entailed by this exercise revealed that some of the most commonly-produced measures were not actually considered to be highly important, while some less-commonly-produced ones were ranked as very important. The importance of a measure derives from what is or could be done with that measure. There is no point in producing a value measure if it does not inform any resulting action. Such actions can include internal efforts to improve value (e.g. a measure of user satisfaction may be used as a benchmark against which to assess

decides to measure ‘value’ using a particular suite of indicators (and assuming that ‘increased value’ is the overall goal), then the collective efforts of the NSO will, rightly or wrongly, aim at enhancing those chosen indicators, perhaps letting other aspects of value that are not included in that suite fall by the wayside. For example, if ‘number of retweets of NSO’s Twitter posts’ is taken as a measure of user engagement, might this inadvertently induce social media managers to post more potentially popular material, maybe at the expense of more ‘serious’ statistical material that would better inform users but be less susceptible to being retweeted? And is this a good thing or a bad thing, or does it not matter? Another example: if error detection time is taken as a measure of accuracy, might this lead to statistical teams taking longer to check and release their figures, and in so doing reduce timeliness? This is not to say that this would be a bad thing—but simply that this possibility, or more generally the possibility that adopting a measure might influence behaviours, should be taken into account.

efforts to increase user satisfaction); or to prove value (e.g. a positive measure of public trust in official statistics can be used as a marketing tool)

4. Distinguishing between production-based and consumer-based concepts of value leads to different approaches to assessing value

87. When we view the value of our products in terms defined by the NSO and its production processes, we end up with measures of value that are also defined in these terms, based around what we make and what we do. If we switch to a consumer-based conceptualization of value, the range of possible measures is different, broader and potentially more helpful in guiding us towards targeted endeavours to improve our value.

5. No single value measurement framework will be suitable for all countries

88. Different levels of organizational maturity and capabilities, different core missions and strategic goals, and different kinds of statistical products and services across countries mean that it would be impossible to produce a single framework of value measures and recommend all countries to produce them all. Cross-country comparability in value measures is therefore not a realistic prospect, nor one to which countries aspire since their attempts to measure value are principally intended for domestic purposes. Nevertheless, the review of countries' experiences did reveal commonalities, allowing for a common direction in broad scope if not in detail. Furthermore, the review revealed that value measures may lack meaning if treated in isolation: they are often more useful when looked at in groups of related measures (such as a series of measures of use of statistics). The measurement framework should not therefore be viewed as a menu or wish list from which to select individual indicators at will, but rather as a guide to dimensions and sub-dimensions on which countries may elect to focus.

B. Recommendations

89. In conjunction with the framework put forward for measuring the value of official statistics, the Task Force recommends that NSOs consider taking the following actions to support such measurement:

1. Start with a goal and build a Results Map

90. While we have proposed here a generic common goal that applies to all countries ('official statistics add increasing value to our communities'), NSOs may wish to adapt this goal to align with their own mission statements or strategies. Building outwards from their goal, NSOs should determine their own mid-term strategies and tactical outcomes, using the Results Map presented here as a model or as inspiration if they wish. While the present work proposes a measurement framework in general terms, the actual measures selected by any given country to measure the value of their official statistics should fit into their own Results Map, with every measure used being justified by a clear connection to the outcomes they are aiming for.

2. Find out from users what is valuable to them

91. Do not assume that we know what constitutes value in the eyes of our users. Seek evidence of the criteria determining value, the relative weights applied to them, and the ways that these differ among users groups and uses. We must acknowledge that asking people 'what makes statistics valuable to you?' is not the best way to do this. More subtle methods of determining people's preferences must be employed, since we are rarely conscious of and able to articulate our own values.

3. Keep asking

92. Conduct research on user perspectives (what criteria they use to determine value, as well as how well those criteria are met by our products) on a continuous basis. User needs are not static, as we have seen from section 0 (pandemic impacts on value), so the methods we use to assess how well those needs are met also cannot be static.

4. Ask everyone

93. We cannot assume that all components of value are equally important to all users. We must segment our user (and passive user and non-user) base and investigate their needs and preferences separately.

5. Consider consequences of a measure becoming a target

94. Pay careful attention to the actions that might be promoted if inclusion of a measure in a value measurement framework leads to efforts to affect that measure. Such unintended consequences could be positive but are often negative.

6. Review and revise measures and methods

95. Deciding upon a suite of measures by which to assess the value of our work is not a one-time undertaking. Users' needs will evolve; our products and services will evolve; the strategic goals, strategies and outcomes of our organizations will evolve. Hence the appropriate set of indicators for measuring value will also change. Available methods will also improve over time. An NSO may wish to refer back to their organization's Results Map to see which aspects have changed and should be updated, rather than undertaking the entire exercise from scratch each time. This kind of review should be integrated into the regular processes of the organization.

7. Make use of the value measures produced

96. Official statisticians love to repeat the adage that statistics are not useful unless they are used. Only then can their benefits be delivered. This applies just as much to measures of our own value as it does to classical statistical measures. We should not produce 'value metrics' merely for the sake of doing so, but should endeavour to act upon the findings, to improve what we do. Relatedly, if we find that a value measure is not routinely used to drive change, we should question its importance and, if we find it not to be justified, discontinue its production.

C. Further work**1. Continue gathering national examples**

97. The case studies included in the current version of this work were collected in tandem with the effort to review the measurement framework. Experiences from countries informed the decisions about retaining or setting aside measures from the framework. Some of the case studies, therefore, refer to experience with measures which, ultimately, have not been retained as potential measures of value. An important exercise will be to gather a wider selection of case studies from across the CES countries, with a focus on those which have been marked as having potential as indicators of value. Countries could make use of the existing agreed case study format to enable them to share experiences and lessons learned.

2. Transform this work, especially national examples, into a living online tool

98. As agreed early on in the current work, and reiterated by the CES Bureau when reviewing an interim progress report in October 2020, the usefulness of national case studies will be greatest if they can be housed in an online format, which can be added to and changed as experiences accumulate. Further work could therefore examine possibilities to develop this entire report into an online environment, with a particular focus on building a living repository of country experiences. This would need to be accompanied by a plan for maintaining such a repository on an ongoing basis.

3. Continue international collaboration to share and improve

99. The wealth of ideas, knowledge and experience among NSOs is immense. The present Task Force has found enormous benefit in regular exchange on everything from high-level conceptual questions (defining value) to detailed methodological discussions (techniques for measuring some of the indicators, techniques for ascertaining user preferences, etc.). The

imperative to ask ourselves difficult questions about why we are doing what we do, whether we are doing it well, how we can improve and how we can demonstrate our worth is common to all countries.

100. It is essential that this does not end with the present report but is extended into the future through continued sharing of experience on the successes and failures of different value measures, methods for determining users' perceptions of value, and ways in which this information is used to inform actions designed to improve the value of official statistics.

101. Possible means of achieving this might include

(a) Organizing international workshops for the presentation of country case studies;

(b) Organizing a workshop to develop a more refined generic form of the Results Map, leading to development of guidance for countries on employing this technique;

(c) Inviting countries to share experience with employing the approaches proposed in this document to develop their own national frameworks for assessing value.

4. Communicate the results of this work

102. Countries that have succeeded in producing measures of value may have individually communicated their results on a national level. It may also be helpful for the official statistics community as a whole to ensure that the collective effort is communicated widely. This would require interested NSOs to champion the cause as well as being done at the international level.
