



Standards Alliance: Phase 2

Annual Report Year I July 12, 2019 to July 11, 2020

Submission Date: October 9, 2020 Version: 2

Agreement Number: 7200AA19CA00012 Activity Start Date and End Date: July 12, 2019 to July 11, 2024 AOR Name: Eleanor Thornton

Submitted by: Leslie McDermott **American National Standards Institute** 1899 L Street NW, 11th Floor, Washington, DC 20036 Tel: 202-331-3626 Email: Imcdermott@ansi.org

This document was produced for review by the United States Agency for International Development (USAID)

Contents

I. PROGRAM OVERVIEW/SUMMARY	3
I.I Program Description/Introduction	3
2. ACTIVITY IMPLEMENTATION PROGRESS	4
2.1 Progress Narrative – Year One	4
Adaptation of project plan due to COVID-19	4
Development Objective: Establish Baseline	5
2.2 Implementation Challenges	7
3. RESULTS ACHIEVED	8
4. SUCCESS STORIES	8
ANNEX I	9
ANNEX 2	10

I. PROGRAM OVERVIEW/SUMMARY

	Standards Alliance: Phase 2
Program Name:	
Activity Start Date And End Date:	July 12, 2019 – July 11, 2024
Name of Prime Implementing Partner:	American National Standards Institute (ANSI)
Agreement Number:	#7200AA19CA00012
Name of Subcontractors/Subawardees:	AdvaMed, Ethical Apparel Africa
Geographic Coverage (cities and or countries)	Brazil, Colombia, Peru, Mexico, Ghana, Kenya, South Africa, Zambia, West Africa (regional), Indo-Pacific (regional)
Reporting Period:	July 12, 2019 – July 11, 2020

1.1 Program Description/Introduction

Standards form the foundation of world trade and the efforts of the Standards Alliance help to create fertile ground for reciprocal trade with the U.S. Through increased adherence and understanding of standards and conformity assessment principles, participating nations will become more competitive in the global market, be more prepared for bilateral trade agreements, and be more capable of protecting their citizens from hazardous goods.

The Standards Alliance Phase 2 (SA2) will engage target populations including, but not limited to, foreign government officials and ministries responsible for standards, trade, and consumer protection; foreign private sector; industry groups; civil society; consumer interest groups; business professionals; trade policy experts; and academia. The objective of this initiative is to build on the past successes, lessons learned and impact measured to-date from the first iteration of the Standards Alliance and to support the capacity of developing countries in the areas of legal and regulatory framework, standards development, conformity assessment procedures, and private sector engagement. Achieving these goals also help countries remove non-tariff barriers, and stimulate economic growth, while also preserving and expanding markets for U.S. companies through more predictable and transparent avenues for trade and investment in partner markets/regions.

The implementing partner for this cooperative agreement is the American National Standards Institute (ANSI). ANSI is a private, non-profit organization that administers and coordinates the U.S. voluntary standards and conformity assessment system. ANSI's mission is to enhance U.S. global competitiveness and the American quality of life by promoting, facilitating, and safeguarding the integrity of the voluntary standardization and conformity assessment system. ANSI's role as a coordinating body and the bridge between the private and public sectors uniquely positions the Institute to build partnerships and foster collaborative solutions that address both national and global priorities. ANSI is also a membership organization, providing members with the broadest access to up-to-date standards policy information and opportunities for participation, leadership, and influence. Finally, ANSI promotes the use of U.S. standards

internationally, advocates U.S. policy and technical positions in international and regional standards organizations, and encourages the adoption of international standards as national standards where they meet the needs of the user community.

The Standards Alliance Phase 2 will include activities in markets representing a variety of geographical regions and levels of economic development, subject to the agreement of USAID. In consultation with USAID Missions, U.S. government, and private sector experts, ANSI will select the countries/regions based on demonstrated commitment and readiness for assistance, as well as U.S. private sector interest and development impact. ANSI expects to engage at both the national and regional level where appropriate and to involve partners in the Latin America, Africa, Middle East/North Africa, and Indo-Pacific regions.

SA2 Focus on Medical Devices to Support COVID-19 Response

In June 2020, USAID recognized the critical role standards and conformity assessment play in advancing public health and safety during global health emergencies and obligated \$3.5 million to SA2, as part of the more than \$1 billion the agency has committed to aid the global COVID-19 pandemic response. This SA2 project will promote regulatory convergence in the context of COVID-19, good regulatory practice (GRP), and the adherence and adoption of international medical device standards. These objectives will also establish an efficient medical device regulatory environment and framework that will facilitate the COVID-19 response and diminish technical barriers to trade—thus promoting trade quality medical devices. AdvaMed—a U.S. based trade association and ANSI member—is the primary private sector partner for this project, which includes select partner countries in Latin America, Africa, and Southeast Asia, and overflowing impacts within those regions.

2. ACTIVITY IMPLEMENTATION PROGRESS

2.1 Progress Narrative – Year One

Since the launch of the SA2 on July 11, 2019, ANSI has worked diligently to ensure high quality management of the program and to optimize support from the private and public sectors—both domestic and international. ANSI successfully submitted key deliverables in the first year of the program, such as the Monitoring, Evaluation and Learning (MEL) plan, implementation plan, needs assessment tools, and the solicitation and selection of proposals.

The SA2 Implementation Plan (dated March 24, 2020) outlined baseline activities to be carried out during the first year of the SA2. This report includes progress made towards those activities, a summary of the outputs, and a description of challenges encountered during implementation.

Adaptation of project plan due to COVID-19:

The spread of the COVID-19 pandemic and the subsequent travel and meeting restrictions have precluded ANSI from hosting in-person events. These restrictions have affected the implementation of SA2 Year I, notably the in-country needs assessments and regional events. However, the SA2 team has increased its capacity to deliver effective virtual events in lieu of in-person meetings. Moving forward, ANSI and its private sector partners will utilize virtual tools to conduct needs assessments and the early stages of many SA2 projects to mitigate health and safety risks related to in-person activities and international travel.

The narrative that follows describes modified activities and adaptations made to the Implementation Plan

in response to the COVID-19 pandemic.

Development Objective: Establish Baseline

1) Selection of Partner Regions/Countries and Conduct Needs Assessments

The preliminary selection of partner countries emerged from discussions ANSI had with both U.S. private sector partners and international partners in countries/regions where USAID works. ANSI shared U.S. private-sector concept ideas with USAID and country/regional partners to gauge potential interest in the proposed activities and refine concepts. ANSI offered guidance to U.S. private sector partners on both past ANSI experience with potential partner countries and country/regional partner experience, highlighting political will, country/sector needs, and ease of doing business. With input from these actors, ANSI finalized its preliminary selection of partner countries/regions outlined in the Year 2 Work Plan. These include Brazil, Colombia, Peru, Mexico, Ghana, Kenya, South Africa, Zambia, West Africa (regional), and Indo-Pacific (regional).

As mentioned above, the pandemic forced ANSI to cancel plans for regional events and in-country needs assessments. However, ANSI began developing a tool to constantly collect information and establish project baselines, as documented in the SA2 Monitoring, Evaluation and Learning (MEL) Plan. Building on efforts from Phase I, the Standards Alliance performed an assessment of existing quality diagnostic tools and guidebooks to inform Standards Alliance programming and to consider best practices for monitoring and evaluating partner country progress. These included guides and diagnostic tools from the Asian Pacific Economic Community (APEC), the Organization for Economic Cooperation and Development (OECD), the United Nations Industrial Development Organization (UNIDO), as well as the World Bank.

Due to the broad-reaching nature of the World Bank Quality Infrastructure (QI) Toolkit report and associated diagnostic tool, the Standards Alliance performed an in depth evaluation of a 275-page report to uncover areas for improvement as well as portions that could be adapted for Standards Alliance purposes. Following this assessment, the Standards Alliance developed a draft QI assessment tool to supplement the existing needs assessment toolkit. Additionally, the Standards Alliance Phase I produced a draft report evaluating the strengths and weaknesses of the World Bank Toolkit to guide future discussions with the World Bank on ways to enhance the existing Toolkit. This report is attached as Annex I.

Based on the initial evaluation report that was prepared by a National Quality Infrastructure (NQI) expert consultant, the ANSI team determined that only a handful of the World Bank toolkit questions were fit for the purposes of the SA needs assessment and would need to be supplemented by tailored questions. The ANSI team worked to compile a list of questions that were more relevant and appropriate for gathering information and data SA 2 will be tracking throughout the program. The ANSI team reviewed the questions, added weighted scores for each question and made modifications to ensure that everyone had the same interpretation and understanding of what was required from the needs assessment tool. The latest version of the SA2 tool is attached as Annex 2. The SA2 team also developed and finalized additional program management, stakeholder management/subawardee management, and risk management. These tools are iterative and will likely be updated during project implementation.

2) Identify Private Sector Partners

To promote the SA2 with the private sector, ANSI hosted an awareness-building meeting in March 2020

to introduce the program to ANSI members and others in the private sector. ANSI followed up with meeting attendees and reached out to additional private sector organizations, based on sectors of focus and geographical regions of interest. The March 2020 meeting and subsequent outreach validated interest from the private sector and resulted in ANSI receiving 20 concept notes for proposed activities. These concept notes were reviewed by ANSI and USAID, and promising concepts were invited to submit a full proposal. As indicated in the Year 2 work plan, ANSI expects to work with, at a minimum, the following private sector partners to execute future SA2 activities:

- Garner Advisors (ECOWAS Clean Renewable Fuels) proposal and agreement in progress
- Center for Water Security Cooperation (WASH-related standards and their reference in law, regulation and policy) concept note approved, proposal in progress
- Regulatory Strategies and Solutions Group (Training in Good Regulatory Practices for Regulatory Impact Analysis (RIA) Teams in Government Ministries) – activity underway September 2020
- ASTM International and the American Petroleum Institute (ECOWAS harmonization of petroleum standards) project proposal submitted and under review
- American Concrete Institute (concrete and building code adoption in Africa) concept note approved, proposal in progress
- IAPMO (Increase the Flow of WASH Services) proposal conditionally approved, scope reduction and agreement in progress
- American Water Works Association (Utility Management Standards Training for water sector utilities) proposal approved, agreement in progress
- NSF International (Community Water Systems Standards for safety and risk management) concept note approved, proposal in progress
- International Code Council (energy efficient building codes in Mexico) concept in development
- Ethical Apparel Africa (COVID-19: Surgical Mask Production Project) agreement in place, activity underway September 2020
- AdvaMed (COVID-19 Medical Device Regulatory Convergence Project) agreement in place, activity underway September 2020

3) Development of Strategic Partnerships

According to the strategic objectives set for the SA2, the SA2 prioritized the exploration of collaboration opportunities with organizations that provide technical assistance on NQI matters, and the development of strategic partnerships as appropriate. This work allows the SA2 to maximize the impact of its technical assistance by leveraging other organizations' resources. It will also ensure that the technical assistance approach of other organizations is balanced in terms of the value that they provide to a system built around voluntary consensus standards and the adoption of a multiple path approach.

With USAID/TRR's support to connect and follow-up with these organizations, ANSI conducted the following activities:

- World Bank: ANSI, USAID and the Office of the U.S. Trade Representative (USTR) met with the World Bank's Markets and Technology group in January 2020. The purpose of the meeting was to introduce the new World Bank lead for the Quality Infrastructure (QI) work, and provide updates on recent work related to QI. Both sides agreed to

continue collaboration around bringing U.S. knowledge into existing country work by the World Bank; as well as updates to the World Bank's QI toolkit. ANSI and the World Bank continued communicating through spring 2020, until ANSI learned that the majority of the World Bank's QI portfolio would be redistributed in another re-organization and possibly closed down. The SA2 will continue to monitor these developments and reassess potential collaboration opportunities as appropriate.

- African Union: Under the Standards Alliance Phase I, USAID and ANSI successfully concluded a partnership agreement with the African Union Commission (AUC) to support implementation of the TBT Chapter of the African Continental Free Trade Area (AfCFTA). As the SAI comes to a close in 2021, ANSI will explore opportunities to continue the partnership with the AUC under the SA2 or another mechanism.

4) Coordination with USAID Bureaus and Mission (and other U.S. government agencies)

Beginning in fall 2019, USAID and ANSI worked together to brief key U.S. government agencies and staff on the Standards Alliance Phase 2, and to discuss opportunities for cooperation where appropriate. Briefings were held with USAID Bureaus and Missions, as well as the interagency committee for technical barriers to trade administered by USTR. During Year I, the SA2 worked closely with USAID partners at country missions, regional trade hubs, and regional programs in the selection of partner countries and review of project proposals. In particular, the Standards Alliance worked with missions in Mexico, Peru, Brazil, Ghana, Cote d'Ivoire, West Africa, and Southern Africa to refine activity proposals for year 2. The most robust engagements are described further below.

- In West Africa, the Standards Alliance coordinated with country missions in Ghana, Côte d'Ivoire, Senegal, and Nigeria to assess a proposal that will support the harmonization of regional petroleum standards. Recognizing the importance of USAID's regional programs, the Standards Alliance requested guidance and feedback on the early draft proposal with the West Africa Trade and Investment Hub as well as the USAID-funded Analytical Support Services and Evaluations for Sustainable Systems (ASSESS) program based in Ghana. Feedback gathered from the multiple levels of USAID programming supported subtle adjustments to this, and other proposals to ensure that they align with, and play a complementary role for ongoing local and regional programming;
- In Mexico, USAID led outreach to the country mission to identify support for SA2 programming. A concurrence memo was developed and, incorporating Mission feedback, the SA2 also engaged with a USAID-funded project to discuss potential energy-related programming opportunities; and
- Regarding the COVID-19 MDRC, USAID connected ANSI to stakeholders that will be important for the future execution of the Medical Devices COVID-19 project, including the USAID Global Health Bureau, and the project staff of the Promoting the Quality of Medicines (PQM+) program.

2.2 Implementation Challenges

As previously described, the COVID-19 pandemic halted travel and in-person events, which was the primary implementation challenge faced in Year 1. These restrictions hindered in-country needs assessments, activities planned pre-COVID-19, and shaped project development for future activities.

Despite this formidable challenge, ANSI adapted its management style to support continued implementation of SA2, and became a partner in USAID's broader efforts supporting the global COVID-19 response by promoting quality standardization, conformity assessment, and technical regulations. The SA2 team used this challenge to enhance the delivery of virtual events to minimize future impediments caused by travel restrictions. With a slightly altered schedule and format, the final stages of the needs assessments as well as other work plan activities will proceed in Year 2.

3. RESULTS ACHIEVED

Annexes I and 2 of this report includes the deliverables that were successfully completed during this program year. Year I deliverables include a report containing the assessment of the World Bank QI toolkit, needs assessment tools, and the solicitation and selection of proposals. Additionally, SA2 successfully concluded an Agreement modification finalizing the obligation of \$3.5 million for COVID-19 related activities. The finalization of the scope of work and work plan for these funds is included in the Year 2 Work Plan. Sub-award agreements were concluded between ANSI and Ethical Apparel Africa, and between ANSI and the AdvaMed. Both of these sub-awards will support execution of the above-mentioned COVID-19 funds.

In the forthcoming Annual Performance Report for the SA2, ANSI will include further updates on results achieved during the 2019-2020 project year, including progress towards deliverables under the COVID-19 MDRC activity, as well as reporting on indicators listed in the SA2 MEL Plan.

4. SUCCESS STORIES

None to report yet, although the SA2 is developing a project fact sheet as well as a flyer for the COVID-19 Medial Device Regulatory Convergence (MDRC) project. These documents as well as any relevant success stories will be included in future reports.



Evaluation of the World Bank QI Toolkit



SA2 Needs Assessment Tool



Evaluation of the World Bank QI Toolkit

17 April 2020, 4:00 pm EAT

Submitted to: American National Standards Institute 1899 L Street, NW, 11th Floor Washington, DC 20036 Tel: 1.202.293.8020 E-mail: David Jankowski, <u>djankowski@ansi.org</u> Jenny Mleczko, <u>jmleczko@ansi.org</u> Via Electronic Submissions

Submitted by:

Jonkheer Consulting Franz Leharlaan 134 2102 GW Heemstede The Netherlands 0031 23 5451630 / 0031 620 192535 E-mail: k.r.jonkheer@jonkheerconsulting.nl



Table of Contents

Executive Summary	.3
, 1.Background World Bank QI Toolkit	.5
2.Characteristics of the World Bank QI Toolkit	.6
2.1. Introduction	.6
2.2. Objective of the Evaluation	.9
3.Evaluation of Generic Usability	.9
4.Evaluation of Functional Usability1	.7
5. World Bank QI Toolkit compared to other relevant QI Documents and Methodologies .2	20
6.Conclusion2	27



Executive Summary

This document provides a comprehensive assessment and qualitative evaluation of the World Bank Quality Infrastructure Toolkit (the Toolkit). The assessment begins with an overview of the Toolkit before breaking down the teasing out the Toolkit's characteristics. This background, lays a foundation for the heart of the evaluation, which focuses on the utility, or usability, of the Toolkit from generic and functional levels of analysis. Generic analysis focuses on the general readability and educational value of the Toolkit, while the functional analysis considers the Toolkit's performative value to inform and/or guide implementation of the elements of Quality Infrastructure (QI). Following this analysis, a comparison of the Toolkit to other QI documents and methodologies is provided. This evaluation aims to inform organizations that hope to use the Toolkit to assess national QI and inform the improvement of an existing QI based on that assessment.

Overall, the Toolkit is a comprehensive tool to aid countries in learning about and considering their own QI. It provides a clear summary and definitions of each element of the QI system. However, the Toolkit remains vague as it relates to the implementation of a QI system. While this is likely to allow countries to implement each element in a manner that best fits local needs, there could be more clarity to support the development of a national implementation plan.

The Toolkit scores well on 'generic usability'; it is an excellent diagnostic tool. Its actionability, coherence, consistency and comprehensiveness, explanation of complex issues as well as the use of Annexes (the Annex is the Comprehensive Diagnostic Tool itself), are effective. While dealing with each QI-topic in a proportionate way, and effective use of tables, maps and figures scores excellent.

However, the Toolkit assumes well-developed institutional environments. Countries with weak institutions may find the Toolkit hard to use due to its elaborateness. For example, the diagnostic may conclude the absence of accreditation in a country, however some countries will never achieve having accreditation facilities, not at least in the near future. In addition, steps following the diagnostic or gap analysis of the QI are unclear. For developing countries, focusing on developing QI-processes may be far more relevant and effective, even better when dedicated around value chains. This is a critical point as developing countries often focus on a limited number of products for international trade, for example beef from Botswana, coffee from Uganda, and bananas from the Caribbean.

The 'functional usability' in terms of reliability, validity and precedent overall scores reasonable. Only the criteria of 'assessment,' defined as the process of collecting and interpreting relevant information about a country's QI, scores excellent. And, the idiographic understanding, an understanding of a particular individual case, scores effective, mainly because the Toolkit covers an exhaustive set of QI-elements serving as a 'mirror' to an individual case. However, the Toolkit scores 'limited' or at best 'sound' on matters of reliability and validity, due to the complexity of the instrument as well as the subjectivity of the assessment of its elements. Additionally, the evaluative portion of the Toolkit is predicated on the availability of a QI expert, while the qualifications of such an expert are left ambiguous opening a door for divergent interpretation depending on the knowledge of the "expert".

Finally, internationally relevant parties like UNIDO and ISO do not endorse the Toolkit, as these parties have indicated that they were not sufficiently consulted in the Toolkit's development.



A brief list of the strengths and weaknesses of the World Bank QI Toolkit are listed below:

Strengths:

- Comprehensive account of QI elements, covering all aspects of the QI-domain;
- Strong educational tool for individuals learning about QI;
- Overall, an excellent diagnostic and effective evaluative tool.

Weaknesses:

- Complex and complicated creating opportunity for confusion in areas without well-developed institutions;
- Does not accurately capture the multiple pathway approach to standards development and reflects a Eurocentric view for some items (ie – definition of international standard and mutual recognition);
- Assessment of the different QI-elements may invite subjectiveness through the abundance of parameters and criteria per element and the vagueness of the requirements for the assessor (the expert);
- Is purely a toolkit consisting of 'a set of tools' and is not 'a methodology';
- Underexposed the QI-process;
- Had limited oversight from relevant international organizations (ISO, IEC, UNIDO);
- Does not address (intended) results of the QI.



1. Background World Bank QI Toolkit

The content of the publication "Ensuring Quality to Gain Access to Global Markets - A Reform Toolkit", accompanied by the "Comprehensive Diagnostic Tool - Annex to the QI Toolkit", in this document referred to as the Toolkit, was developed jointly by the World Bank Group and the National Metrology Institute of Germany (PTB). The Investment and Competition Unit of the World Bank's Macroeconomics, Trade and Investment (MTI), initiated the preparation of this toolkit, by Christine Zhenwei Qiang, practice manager for the Global MTI, Global Practice. Catherine Masinde, practice manager for the MTI Global Practice's Global Business Regulation Unit, subsequently led the development of the Toolkit. The principal author of the Toolkit is global quality infrastructure (QI) consultant Martin Kellermann.

Using their vast experience in upgrading and reforming QI ecosystems, the World Bank and the National Metrology Institute of Germany (PTB) have partnered to develop the first comprehensive QI diagnostic and reform toolkit, which is **designed to help development partners and country governments analyze their QI ecosystems and develop a coherent offering to support QI reforms and capacity development**. Based on the results, the Toolkit provides recommendations to bridge gaps in the QI ecosystem, support reforms, and build the capacity of institutions (see Figure 1). The Toolkit is a valuable knowledge base for other interested parties to learn about QI and reform all or parts of their QI systems. Such reforms could focus on one, or any combination of, the following objectives:

- Improving the legal and institutional framework for efficient and effective QI
- Enhancing trade opportunities by removing unnecessary nontariff barriers and technical barriers to trade through harmonization of technical regulations and mutual recognition of conformity assessments
- Integrating into global value chains
- Enhancing overall quality of products and services
- Encouraging innovative products to be entered into high-value-added markets
- Increasing productivity and efficient use of scarce resources
- Providing for greater consumer protection

Figure 1. The Toolkit workflow: Reforming the Quality Infrastructure





The Toolkit consists of 12 modules to provide a valuable knowledge resource as a holistic reference - supported by practical case studies and examples - for QI diagnostics, reform interventions and approaches, and monitoring and evaluation.

The Toolkit is peer reviewed by relevant international organizations - namely, the International Bureau of Weights and Measures (BIPM), International Accreditation Forum (IAF), International Laboratory Accreditation Cooperation (ILAC), and International Organization of Legal Metrology (OIML).

2. Characteristics of the World Bank QI Toolkit

2.1. Introduction

The World Bank QI Toolkit follows a specific logic, starting from the policy and legal environment before dealing with each of the QI-elements. The outcome of the evaluation provides qualitative results that an expert can turn into quantitative results. Beyond in-depth reports, the results can be articulated in dashboard-type images for a more rapid understanding of situations when discussing them with counterparts.

This assessment focuses the elements of the Toolkit that may, and may not, be useful. For its use in the assessment of the National Quality Infrastructure of Côte d'Ivoire, the Toolkit should be more of an aid in the elements and items that are being assessed, rather than the basis.

It is important to note that the Toolkit is not endorsed by UNIDO, ISO and others. These parties felt they were not consulted properly during the Toolkit's development. Despite this, there is a lot of useful information in the two documents.

Pillars

The "pillar and building block" approach consists of constructing a diagnostic tool for each identified element of QI. The "effectiveness" of each QI element is considered in relation to four pillars:

- Pillar 1: Legal and institutional framework, in which the broader environment, within which, the entity is legally established and operating is considered.
- Pillar 2: Administration and infrastructure, in which the organizational structure and the necessary infrastructure of the entity to fulfill its responsibilities are considered.
- Pillar 3: Service delivery and technical competency, in which the output and services of the entity are considered, with special emphasis on their demonstrable quality.
- Pillar 4: External relations and recognition, in which the liaisons of the entity with relevant regional and international organizations are considered in view of the need to be acknowledged for its output and services.

Each pillar consists of building blocks that have to be in place for the QI element to function optimally and to comply with international best practices. Some of the building blocks for each of the QI elements would be similar, but there will also be quite a few differences between pillars.



Building blocks

The Comprehensive Diagnostic Tool endeavors to provide a qualitative and quantitative approach for each QI element, which can be visualized as a "*building*" showing the state of implementation through different-colored "bricks", a radar-type diagram for the individual elements, or a dashboard illustration for the QI collectively, supported by an extensive narrative. For each building block, the diagnostic:

- Provides details about the best practices against which the building block should be compared, under the heading "What is meant";
- Shows how the building block can be demonstrated (that is, describing the elements that indicate that the practice exists), under the heading "How can it be demonstrated"; and
- Shows where the assessor could find information to support the existence of such practices, under the heading "Existing information/reporting/ monitoring."

Classification Status

In allocating a quantitative measure to the various building blocks, the Toolkit considers the weight of each building block to clarify how fundamental a specific QI element is to a nation's QI. The most foundational elements are considered "fundamental." At a second level are the "major" building blocks: those necessary for the service delivery to be effective and efficient. At the third level are the "minor" building blocks: those in which the custom and practice of the country play a role rather than international practices. The quantitative evaluation has to recognize these differences.

Consequently, for each building block, an indication as to whether it is "fundamental," "major," or "minor" is provided. This will help the assessor determine the extent and significance of the gap between the current situation and international best practices. In turn, this will be an indication of the "effectiveness" of the QI elements in the country, leading to a judgment call on how much support the country would need to develop its QI to the point where it meets the needs of its stakeholders.

The Toolkit proceeds by separating elements into three broad categories based on relevance for specific development contexts. These include: (a) the basic QI (relevant for a low- or middle-income country approach); (b) an advanced QI (relevant for an economy-wide approach); or (c) ultimately as a mature or innovative QI (relevant for a high-income economy or world-class approach).

Implementation Status

Therefore, the Comprehensive Diagnostic Tool endeavors to provide a qualitative and quantitative approach for each of the QI elements, which can be visualized as a "building" showing the state of implementation through color-coded "bricks", a radar-type diagram for the individual elements, or a dashboard illustration for the QI collectively, supported by extensive narrative.

The evaluation contains a complex array of levels of (a) implementation ("implemented," "mostly implemented," "partially implemented," or "not implemented"); and (b) classification ("fundamental," "major," or "minor"). A judgment call will need to make to determine how far a project wishes to take the capacity-building exercise. A reasonable approach would be to first, address the "fundamentals" and "major" issues. The "minor" issues are, to some extent, "nice-to-haves" or "non-mandatory" and would be included resources permitting. To depict the construct the visualizations, the implementation status of each building block has to be given a numerical value (percentage implemented).



In the Comprehensive Diagnostic Tool, the expert assessing the QI will have to provide a quantitative and qualitative result based on the expert's experience. This includes a narrative in the various sections and must be an evaluation based on a matrix-type approach. The question-and-answer methodology in the Rapid Diagnostic Tool (discussed in module 9, section 9.1, of the Toolkit) provides some guidance in this respect. Once the percentages are determined, it is easy to construct a radar diagram; a radar diagram is drawn on the QI entity's implementation status. To depict the "building" is on the next step.

The percentages can be grouped into four categories, such as the following:

- More than 75.1 percent: Implemented
- 50.1–75 percent: Mostly implemented
- 25.1–50 percent: Partially implemented
- 0–25 percent: Not implemented

The four groups (or more, if the four are considered too coarse a grading) can then be given different colors in the "building". It helps if the colors are chosen to coincide with a color scheme that can be psychologically understood by the potential readers.

Assessment of National Policy and Legal Environment

Two crosscutting issues that have a distinct influence on the quality infrastructure (QI) landscape in a country are the policy environment regarding the QI and the technical regulation regime. These are most commonly contained in a national quality policy and a technical regulation framework. Of each, 1) Benchmark and Significance & 2) Classification, best practices, and implementation strategy should be assessed, in the same way the building blocks are assessed (see above):

- "What is meant"
- "How can it be demonstrated" and
- "Existing information/reporting/monitoring"

The need for expert knowledge

The annex of the Toolkit's comprehensive evaluation criteria emphasizes assessments "do not negate the necessity of expert knowledge" when conducting a comprehensive evaluation of a country's QI. The criteria are primarily guidelines aiming to ensure all important issues are included in an evaluation. The difference between the country's QI and (a) its compliance with stated or formal criteria; and (b) the criteria for competent, effective, and efficient working structures are important to highlight during an evaluation. The former can be provided on paper as a checklist to be ticked off, but the latter depends on the judgment, and hence experience, of the evaluator, as well as quantitative evidence. These are not easy to encompass in a publication such as this.

The requirements for 'an expert', or sometimes 'assessor,' remain ambiguous. The profile of such an expert is also unclear: should it be an intermediate expert, technical expert, an independent expert, etc.



Review Standardization

There is no single approach to implementing the Technical Barriers to Trade (TBT) Agreement. However, twenty years of discussion within the TBT Committee has resulted in the development of guidance that aims to help World Trade Organization (WTO) members find ways of improving how they use the TBT Agreement. To provide guidance on the development of international standards, the WTO TBT Committee determined a set of **principles**: (a) transparency, (b) openness, (c) impartiality and consensus, (d) effectiveness and relevance, (e) coherence, and (f) development dimension (WTO 2000). Organizations, such as the International Organization for Standardization (ISO) have added three more principles to this mix (ISO 2010): (g) stakeholder engagement, (h) due process, and (i) national implementation. Although these nine principles were established for international standardization, they are relevant for regional and national standardization, and form the basis of good standardization practice (GSP). The Toolkit has adopted these principles in module 3, section 3.4.

The **level of maturity** of the country's trade, technical regulation regime, industrial development, and other factors influence the maturity demanded of the national standards body (NSB). A four-level breakdown is shown in table 3.1 of the Toolkit. Levels of maturity must be taken into consideration when a comprehensive diagnostic evaluation of standardization, as well as any other QI-topic, is conducted, thereby influencing the qualitative and, especially, quantitative outcome of the application of the various QI building blocks.

2.2. Objective of the Evaluation

This evaluation aims to provide insight on the generic and functional usability of the Toolkit. The scope of this assessment is not to have reviewed the Toolkit in its entirety, but, rather, to evaluate strengths and weaknesses and to consider potential uses as a guide to inform QI development. USAID and ANSI wish to understand which elements are most valuable, and which should be improved upon to support consistent and effective QI development implementation.

3. Evaluation of Generic Usability

Generic usability refers to how the Toolkit 'communicates' with the user, this could also be considered the general educational value of the Toolkit. This assessment is based upon a set of 12 criteria that are derived from an analysis of several publicly available Review Checklists, used for reviewing Guideline documents, Manuals and/or Toolkits, as well as templates for Review documents (for example: the EU document on Impact Assessment Reviews).

The following 12 relevant criteria for review of the Toolkit have been identified:

Criteria for Assessing Generic Usability of Guideline documents, Manuals and Toolkits					
1	Audience				
2	Is the Toolkit Described in an Actionable Manner?				
3	Clear Document Structure				

JONKHEER consulting

4	Coherent
5	Contain a Good Description of Methods
6	Consistent, Comprehensive
7	Read as a Single Document with Appropriate Cross-Referencing
8	Make Effective Use of Diagrams, Tables, Figures to Support the Text
9	A Clear Explanation of Complex Issues
10	Cover Each Toolkit-topic in a Way that is Proportionate to its Importance
11	Supporting Examples and Good Practices
12	Annexes with Practical, Additional Information that Support the Text

1) Audience

The document has two forewords, from the World Bank and from PTB. The World Bank's foreword states, "This guide is designed to help development partners and governments assess and analyze a country's QI ecosystem; identify issues and gaps; and provide recommendations for how to bridge those gaps and build institutional capacities".

Subsequently, the World Bank's Foreword identifies 'countries' as their audience: "This diagnostic is critical for understanding and identifying gaps and shortfalls quickly, so countries can effectively identify areas for reform. QI is, therefore, a relevant ingredient for achieving the World Bank Group's twin goals of ending extreme poverty and promoting shared prosperity by the end of 2030 through competitiveness, trade, health and safety, and so on. The Toolkit provides a useful framework for helping countries understand where and how to begin the reform process".

PTB's foreword states, "...the German government is naturally committed to enabling its partners in emerging and developing countries to access new markets and strengthen their competitiveness by enhancing the quality of their products. In the framework of our technical cooperation, we place special emphasis on the core of our own quality production: a well-functioning quality infrastructure (QI)".

The mentioned audiences of the two Forewords are not necessarily the same, and although this Foreword refers to an audience, the document itself mostly does not directly refer to a reader. Instead, it remains rather a static, technical review of QI-topics and issues.

Status of document is not clear

It is important to note, that the status of the document is not very clear. The abovementioned section of the World Bank's Foreword refers to itself as a 'Guide'. However, the title refers to a 'Reform



Toolkit'. The document is later referred to a 'Diagnostic Tool' that uses a 'systematic Methodology'. In short, it is not clear what the writers have intended for this document: is it a Guide?, a Toolkit?, a Diagnostic Tool?, a Methodology? The status of the document has its consequences for its audience. Whether this document is a Guide, a Toolkit, a Diagnostic Tool, or a Methodology changes how and by whom this document is, or should be used.

2) Actionable

In principle, the Toolkit may be called *actionable*. For example, Part 3 of the document, Chapter 10 "How to Reform: Interventions and Approaches", is clearly quite actionable: review, design, develop, reform reengineering, establish, are all actionable terms bringing forward practical recommendations. However, all of these elements are described in '*as is*' or '*should be*' modes, and do not necessarily provide '*how to*' steps of a process, resulting in '*how things should be*'.

As an illustration, we consider this small section from Chapter 10:

"Things get a bit more complicated if the institutional framework has to be reengineered. A typical example would be if the NSB loses its mandate to develop and implement technical regulations or mandatory or compulsory standards. In this case, the responsibilities and activities of the NSB must be transferred to a regulatory authority—either an existing one or a new authority that must be established. This will require a new set of legislation to start the process. The actual transfer will need to be carefully planned to ensure that the transitional period does not lead to an "anything goes" situation in the marketplace."¹

This example reads like a set of requirements, however does not read as an offered, proposed practical step-by-step 'how to' solution. In that sense, sections like these (and, the document has chosen this method of description as the predominant style) do not read like a Toolkit, as in 'what to do', 'how and when to do it', and/or 'how to get there'.

A separate example comes from Chapter 9 "Diagnostic Tools": "The various elements of the QI are interrelated, and coordination of their responsibilities and services is an important parameter. Hence, while dealing with the various elements of the QI individually, their overall coordination should not be neglected.²" How the various elements of the QI may be interrelated and how coordination of responsibilities and services should be taken on, remains unclear.

Use of Actionable Terms is not the Same as Guidance on How to Take Action

This can be illustrated with a practical case. In 2016, Egypt drafted a new Metrology Law with the aim to address the lack of coordination within the metrology domain and more specifically activities concerning market surveillance activities of the country, in terms of reliability of use of measurement instruments. The Law proposed the introduction of a new layer of institutions on top of the existing infrastructure of institutions. This new layer would consist of, among others: a Metrology Council, a National Metrology Institute, a National Legal Metrology Institute, and a Central Metrology Authority.

At this stage, it was necessary to consider whether this new layer of institutions would effectively tackle the existing coordination issues. As an example, the so-called Measurement Instruments Life Cycle was analyzed from Manufacturer, Place on the Market, Market Surveillance (Authorities), to

¹ See: See Chapter 10 Implementation, 10.2.2 Reengineering the institutional framework.

² See: Chapter 9 Diagnostic Tools, 9.2.2. Approach of the Comprehensive Diagnostic Tool.



actual Routine Inspections. It was necessary to assess whether the proposed solution, added a superfluous layer of institutions to existing agencies. Additionally, was the new Law necessary, or would 'lower regulation' be sufficient to address the identified issues. The conclusion of the deliberations was that a new layer of institutions would not improve coordination of parties. In fact, it was likely this could further complicate the system. Rather, the coordination between existing institutions should be improved by introducing a set of practical guidelines. As an alternative to this solution, a replacement of institutions was considered (especially, when this would mean better alignment with international best practice).

This demonstrates that it is one thing to identify a certain issue, even if done in 'actionable' terms, but it is another to address issues in terms of how to solve and act upon a problem.

First and Foremost a Diagnostic Tool

Having observed such, the document seems to be, primarily, a Diagnostic Tool (see Chapter 9), and this diagnostic tool is expressly meant to assess the QI-situation in a country 'as is', or give the theory of 'how things should be'. Regarded as such, however informative and comprehensive, it is a rather static document. Further, the term "Reform Toolkit" suggests this is a 'box of tools' the reader or user may take from the toolkit whatever the user deems suitable in a particular situation. In that sense, a 'toolkit' suggests something separate from a 'methodology,' which describes methods consisting of 'a particular procedure for accomplishing or approaching something'. With regard to the limited reference to its audience, QI-topics and issues are sometimes described with concrete actions. However, the "when to" and "how to" questions, instructive as actionable guidance are barely asked. When and how to use a certain method or technique is essential for an effective Toolkit.

3) Clear Document Structure

The Toolkit has a clear document structure. This is a crucial point to ascertain if this is a toolkit or a 'box of tools', rather than a process methodology. The depicted Toolkit workflow (Figure 1) suggests this is a process methodology, however the chapter-structure does not contain a workflow.

Secondly, apart from the stated objective 'improving the legal and institutional framework for efficient and effective QI', the Toolkit does not make clear how the tool addresses (or feeds into) the other mentioned objectives of a QI. Those elements being: the enhancing trade opportunities by removing unnecessary nontariff barriers and technical barriers to trade through harmonization of technical regulations and mutual recognition of conformity assessments, integrating into global value chains, enhancing overall quality of products and services, encouraging innovative products to be entered into high-value-added markets, increasing productivity and efficient use of scarce resources, and providing for greater consumer protection.

Thirdly, the structure of the Toolkit document and the Comprehensive Diagnostic Tool do not fully correspond. This may create confusion. The Reform Toolkit describes in chapters: 'Standards', 'Metrology', 'Accreditation', 'Conformity Assessment', and 'Technical Regulations', whereas the Comprehensive Diagnostic Tool consists of the following Chapter-structure: 'Standards', 'Metrology', 'Accreditation', 'Inspection', 'Testing', 'Product Certification', 'System Certification', 'Technical Regulation' and 'Legal Metrology'. Although it may seem obvious that topics like inspection, testing and certification belong to Conformity Assessment, correspondence between the Reform Toolkit and the Comprehensive Diagnostic Tool could be more aligned.



4) Coherency

The Toolkit can be considered a coherent document, outlining and explaining the main issues and stages of QI in a logical sequence. Cross-referencing between the Reform Toolkit document and the Comprehensive Diagnostic Toolkit may be challenging due to different chapter structures.

5) Good description of methods

Excellent Tool for Diagnostic and Evaluation

The Toolkit describes each element of a QI in great detail and is quite comprehensive. It is an excellent tool to assess the country's current state of QI. In that sense, the diagnostic and evaluative properties of the Toolkit clearly describe the composition of pillars and building blocks for each QI-topic. The Toolkit consists of specific building blocks allowing each QI-topic to be addressed quite distinctively.

As stated in Chapter 2, the Toolkit follows a specific logic, starting from the policy and legal environment before dealing with each of the QI-elements. The outcome of the evaluation provides qualitative results an expert can turn into quantitative results. Over and above in-depth reports, the results visualization tools and dashboard-type images that support a more rapid understanding of situations for discussions with counterparts.

Possibly Complicated to Apply to the Full

For an average professional user or reader, the large number of parameters/components may pose a challenge. Many of these parameters/components are highly dependent on a rather subjective 'expert knowledge'. The methodology included no less than ten parameters/components: Pillars, Building blocks, Classification Status, Implementation Status, Assessment of National Policy and Legal Environment (with its own three parameters), Criteria (used in the expert assessment), Principles, and Level of Maturity. This may invite quite different results when using this Toolkit in different circumstances/countries or when applied by different experts.

The World Bank QI Tool is a Toolkit, not a Methodology

The Toolkit seems to be, primarily, a 'toolkit' in the true sense of the word. It is not a systematic methodology or guideline with sequential steps, guiding the user from start to finish. It, therefore, does not necessarily address the fabric of a country's QI-system in which parties need to work in concert with each other. This weakness is key as the QI-process and, in practice, may turn out to be very important for overall effectiveness and to reach its objectives. Three examples of how this weakness could result in major operational issues for a national QI are included below.

First, it is crucial for the mandate of a Standards Body to note whether it may operate at borders in collaboration with Customs to inspect or certify products pre-market, or whether it may only be engaged in market surveillance, only ascertaining its mandate when the product already has passed Customs and is on the country's market. An issue as subtle as this may prove critical to the perception and operation of a country's QI landscape.

Second, some countries may lack the threshold to develop each QI-topic, or function, to full capacity. The Toolkit partly acknowledges this. For example, in addressing Accreditation it recognizes a country may have to resort to a Regional Accreditation Body; however, it does not address the situation when a country may have to combine functions like certification, accreditation, development of standards,



enforcing technical regulations, etc. on a national level by one single standards body. In such a situation, it is crucial that 'legal firewalls' are maintained for an orderly, maintainable and credible QI.

Third, while it is important that regulations are harmonized at an international level, the national level regulatory framework of a QI needs to be coherent. This means, that coordination regarding the regulations on the separate QI-topics is essential. This also goes for implementation. For example, often there are many parties involved in market surveillance, and an effective coordination of tasks is crucial for the functioning of an effective QI.

These examples, all regarding process-like properties of a QI, are not fully addressed in the Toolkit.

Other mentioned Methodologies

When it comes to mentioning methodologies that form the basis of the identified QI-steps (the building blocks), the Toolkit is always clear '*what*' the building block is or what needs to be done. However, '*how*' it should be done is rarely elaborated. For example, with regard to Technical Regulations the Toolkit mentions a Regulatory Impact Assessment should be completed. How, and when in the process, RIA should be done is not made clear. Another example is engagement of stakeholders. It is noted that stakeholder engagement is required, however 'how' and 'when' remains unclear.

6) Consistency, comprehensiveness

Consistency

A strong point of the Toolkit is that its approach has highly consistent elements. This helps the reader or user 'break down' its contents, despite its volume and comprehensiveness. The structure per QItopic across the four Pillars is always the same. In addition, the associated Building Blocks seem to have kept, where possible, a similar structure.

As mentioned under 'Coherency', cross-referencing between the Reform Toolkit document and the Comprehensive Diagnostic Toolkit may be challenging due to differing chapter structures, and, with this, a different identification of QI-topics between the Reform Toolkit document on the one hand, and the Comprehensive Diagnostic Tool document on the other.

Comprehensiveness

The Toolkit describes each element of a QI in detail and is quite comprehensive. As stated in Chapter 1: based on the results, the Toolkit provides recommendations to bridge gaps in the QI ecosystem, support reforms, and build the capacity of institutions. The Toolkit is also a valuable educational resource for parties interested in learning more about QI and reform to their QI systems. Such reforms could focus on one, or any combination of, the following objectives:

- Improving the legal and institutional framework for efficient and effective QI.
- Enhancing trade opportunities by removing unnecessary nontariff barriers and technical barriers to trade through harmonization of technical regulations and mutual recognition of conformity assessments.
- Integrating into global value chains.
- Enhancing overall quality of products and services.
- Encouraging innovative products to be entered into high-value-added markets.
- Increasing productivity and efficient use of scarce resources.



• Providing for greater consumer protection.

In this sense, the Toolkit runs the risk of being overly comprehensive, attempting to address too broad a set of objectives.

7) Read as one document, effective cross-referencing

The Toolkit has a tendency to sum up, categorize and put into order the different topics ('topics' not necessarily read as 'steps' as in a process) of QI.

The methodological connection, reference and relationship between these different topics is clear, and they are always approached in the same manner (pillars, building blocks). No sequential relation between the topics is made however. As stated earlier, the document is a toolbox and not a methodology or guideline in terms of a process. This, while one might suggest that there likely is a logical, sequential relationship between at least some of the different QI-topics. For example, the development of Standards, and Technical Regulations comes before, the development of Certification, and Market Surveillance.

8) Make Effective Use of Diagrams, Tables, and Figures to Support the Text

The Toolkit makes ample use of diagrams, tables and figures to support the text. The main report (253 pages) has no less than 49 Figures, 3 Maps, and 21 Tables. The Comprehensive Diagnostic Tool (209 pages) has 24 Figures and 23 Tables. The Figures, Maps, and Tables provide a strong schematic depiction of applicable contents and are, where possible, consistent in form and shape.

9) Clear explanation of complex issues

The Toolkit has a very clear structure to 'tackle' the different QI-topics (Standards, Technical Regulations, Accreditation, Metrology, etc.), pillars, and building blocks. This means, each QI-topic is approached in the same fashion and is broken down in an easily understandable way for users.

10) Each QI-topic is covered proportionate to its importance

As already stated, the current Toolkit can be considered a coherent document, outlining the main QItopics in detail. The Toolkit makes clear what is needed to arrive at a full development for each QI topic. It does so through consistent recognizable elements, the four pillars, and the building blocks that are mostly, though not always the same due to specific topic characteristics and requirements, similar for each QI-topic. This results in each QI-topic being dealt with proportionate to its importance.

11) Examples and Good Practices

The Toolkit provides evidence of good practice via the concise sections per QI-topic. In the new revised Guidelines these examples and the good practice sections remain included.

That said the Toolkit seems to consist of many parameters/components. Many are highly dependent on what is, subjectively referred to as 'expert knowledge'. The methodology includes no less than ten parameters/components: Pillars, Building blocks, Classification Status, Implementation Status, Assessment of National Policy and Legal Environment with its own three parameters, Criteria (which used in the expert assessment), Principles, and Level of Maturity. This may be confusing, especially



when this Toolkit should be applied in its entirety, as it might invite rather different results when using this Toolkit in different circumstances/countries.

12) Annexes with Practical, Additional Information

The main Annex to the Toolkit is the Comprehensive Diagnostic Tool. Due to its comprehensiveness, detail and consistent structure, it is quite usable in identifying the status and gaps of a country's QI.

Scores

Based on this review of the World Bank QI Toolkit as illustrated above, each of the 12 identified criteria has been scored on a 5-point scale, indicating the generic usability represented in the Toolkit, ranging from Elementary to Excellent (Table 1).

	Review Criteria: World Bank QI Toolkit	Elementary	Limited	Sound	Effective	Excellent
1	Audience		х			
2	Are the Guidelines Described in an Actionable Manner?				х	
3	Clear Document Structure			х		
4	Coherence				х	
5	Contain a Good Description of Methods			X1		
6	Consistency, Comprehensiveness				Х	
7	Read as a Single Document with Appropriate Cross-Referencing			х		
8	Make Effective Use of Diagrams, Tables, Figures to Support the Text					х
9	A Clear Explanation of Complex Issues				х	
10	Cover Each QI-topic in a Way that is Proportionate to its Importance					х
11	Supporting Examples and Good Practices	х				
12	Annexes with Practical, Additional Information that Support the Text				х	

Table 1. Scores of Review Criteria on the World Bank QI Toolkit.

¹This score is a mean of extremes. The World Bank QI Tool is excellent as a diagnostic or evaluative tool, and as such seems quite suitable to assess a nation's QI; however, it is elementary when it comes to describing, let alone explaining, methodologies.

The Toolkit scores 'Excellent' on the use of Diagrams, Tables, and Figures. Due to a complete lack of Examples and Good Practices, this measure has received an Elementary score. Despite many positive scores in this assessment methodology, much remains to be improved. This includes recognizing the audience, partially due to the inconclusiveness about the status of the document. Actionability,



Coherence, Consistency/Comprehensiveness, Explanation of Complex issues and Annexes are all considered Effective.

Remaining issues

Following this review, one unresolved issue remains. This is the question of whether the Toolkit would be more effective if it did not only focused on the rather static **"what's"** of a QI, but also focus on **process and results**. For that, a more LogFrame approach would be worth to consider:

 Log Frame. 'Logical Framework', or 'log frame', describes both a general approach to project or program planning, monitoring and evaluation, and – in the form of a 'log frame matrix' – a discrete planning and monitoring tool for projects and programs. Log frame matrices are developed during project/program design and appraisal stages, and are subsequently updated throughout implementation while remaining an essential resource for ex-post evaluation. Applied to a dynamic process, a log frame may be perceived as rigid in its approach, as well it being potentially incompatible with participatory assessments (wherein the priorities of vulnerable groups can often be overlooked by the actors or dominant stakeholders who take part in the log frame construction).

Because the Toolkit is as much a diagnostic tool as an evaluative tool, a Log Frame could be applied as a stage of Monitoring and Review, focusing on process and results. As an addition to the Toolkit, a Log Frame is suggested as optional because it supports practical guidance in identifying or reiterating:

- The overall project/program goal;
- The purpose: the project's central purpose or outcome a tailored result that the intervention seeks to achieve in support of the overall goal;
- The outputs: those observable, measurable change, and tangible products/services to be delivered by the intervention, which serve to achieve the overall goal and purpose;
- The activities: supporting activities i.e., the main tasks that need to be completed in order for the output to be achieved are defined; and
- The inputs: required inputs to facilitate the process and to achieve the results (goals).

4. Evaluation of Functional Usability

Functional Usability of the World Bank QI Toolkit

The functional usability characteristics of an assessment tool usually focus on types of reliability, validity, and precedent. These components describe how consistently the Toolkit will produce unambiguous results through repeated use:

- 1. Reliability: a measure of the consistency of test or research results.
- 2. Validity: a measure of the accuracy of the results of a test or study.
- 3. Test-retest reliability: a kind of reliability. if the test yields similar results every time it is administered to people.
- 4. Inter-rater reliability: also known as inter-judge reliability. If different judges independently, agree on how to score and interpret a test. It is the degree of agreement among raters. It is a score of how much homogeneity or consensus exists in the ratings given by various judges.
- 5. Face validity: a given assessment tool may appear valid simply because it makes sense and seems reasonable.
- 6. Predictive validity: a tool's ability to predict future characteristics or behavior.



- 7. Concurrent validity: the degree to which the measures gathered from one tool agree with the measures gathered from other assessment techniques.
- 8. Idiographic understanding: an understanding of a particular individual case.
- 9. Assessment: the process of collecting and interpreting relevant information about a client or research participant.
- 10. Standardization: the process in which a test is administered to a large group of people whose performance then serves as a standard or norm against which any individual's score can be measured.

As stated, the Toolkit is quite comprehensive in its approach. In this sense, it contains a lot of valuable information and allows for a very effective diagnostic or gap analysis of a country's QI. Its level of detail is very clear, since the users of the Toolkit will measure the same items to the same degree of detail.

The Complexity of the World Bank QI Toolkit May Affect Its Evaluative Accuracy

However, the Toolkit consists of many parameters/components, many of which are highly dependent on what is, subjectively referred to as 'expert knowledge'. The Toolkit and its proposed methodologies include no less than ten parameters/components: Pillars, Building blocks, Classification Status, Implementation Status, Assessment of National Policy and Legal Environment with its own three parameters, Criteria (which used in the expert assessment), Principles, and Level of Maturity. This may invite a rather strong subjectivity into the application of the Toolkit, potentially yielding rather different results when using this Toolkit in different circumstances/countries. The Toolkit requires expert assessments on many levels and not only for autonomous elements or building blocks, but also in an interrelated fashion. Therefore, a minor or single-element inaccuracy may create a large inaccuracy when assessed along multiple criteria all dependent on expert assessments.

The evaluative accuracy and consistency probably would have been better served if the Toolkit were simplified. For example, if a Diagnostic or Evaluative Toolkit should yield predictable and accurate results, a Toolkit would be better served with 'yes'/'no' questions. In other words, the Toolkit would be more accurate with answers to issues that are simple and limited, unequivocal, and one-dimensional. The simpler, the more accurate. Further, the usefulness of the Toolkit may be undermined by its comprehensiveness. For example, the diagnostic may conclude the absence of accreditation in a certain country; however, some countries will never achieve accreditation, at least not in the near future, which means this result of the diagnostic may remain of limited value.

The Toolkit itself acknowledges its complexity stating, "A comprehensive assessment of the QI of a country is a **complex undertaking**. It is virtually impossible to reduce the outcome of such an assessment to a single figure or a simple pronouncement. There are just too many possibilities and nuances that have to be considered, too many externalities that have an influence.³ (...) The evaluation is therefore a **complex array of levels** of (a) implementation (implemented, mostly implemented, partially implemented, or not implemented); and (b) classification (fundamental, major, or minor). A judgment call is necessary to determine how far a project wishes to take the capacity-building exercise. A reasonable approach would be that the "fundamentals" and the "major" issues must be dealt with first. The "minor" issues are, to some extent, "nice-to-haves" or "non-mandatory," and would be included, resources permitting." This comes on top of making an assessment of the break-down of each building block in *'what is meant'*, *'how it can be demonstrated*, and what *'existing*

³ Chapter 9 Diagnostic Tools, 9.2.2. Approach of the Comprehensive Diagnostic Tool.



information/reporting/monitoring' sources are available, or in other words: how to satisfactory verify the level of development of (each element of) each building block.

Scores

Based on this review of the Toolkit for its functional usability, each of the 10 identified types of reliability, validity, and precedent has been scored on a 5-point scale. Indicating the functional usability represented in the Toolkit, ranging from Elementary to Excellent (Table 2).

	Review Criteria: World Bank QI Toolkit	Elementary	Limited	Sound	Effective	Excellent
1	<u>Reliability</u> : a measure of the consistency of test or research results.		х			
2	<u>Validity</u> : a measure of the accuracy of the results of a test or study.			х		
3	<u>Test-retest reliability</u> : if the test yields similar results every time it is administered		х			
4	Inter-rater reliability (or inter-judge reliability): If different judges independently, agree on how to score and interpret a test. The degree of agreement among raters.		x			
5	Face validity: a given assessment tool may appear valid simply because it makes sense and seems reasonable.			x		
6	<u>Predictive validity</u> : a tool's ability to predict future characteristics or behavior.		х			
7	<u>Concurrent validity</u> : the degree to which the measures gathered from one tool agree with the measures gathered from other assessment techniques.			x		
8	<u>Idiographic understanding</u> : an understanding of a particular individual case.				х	
9	<u>Assessment</u> : the process of collecting and interpreting relevant information about national QI.					Х
10	<u>Standardization</u> : the process by which a test is administered to a large group of people whose performance serves as a standard against which any individual's score can be measured.			x		

Table 2. Scores of Different Types of Reliability, Validity, and Precedent on the Toolkit.

The overall 'functional usability' scores reasonable. Only the criteria of 'assessment' - defined as the process of collecting and interpreting relevant information about a country's QI, scores excellent. And, the idiographic understanding - an understanding of a particular individual case, scores effective, mainly because the Toolkit covers an exhaustive set of QI-elements serving as a 'mirror' to an individual



case. However, the Toolkit scores 'limited' or at best 'sound' on matters of reliability and validity, mainly because of the complexity of the instrument as well as the unavoidable subjectivity of the assessment of its elements.

Level of Usability is Partly Dependent on What the Intended Use of the Diagnostic/Evaluation is

The clarity of the functional usability of the Toolkit is also relevant, with regard to the intended results. Why are countries conducting a QI assessment with the aid of this Toolkit? Is it to identify gaps and allocate resources to fill these gaps? Is it because countries wish to use this assessment to apply for funding, either nationally or internationally (or, both)? To what extent may the assessment be politicized?

5. World Bank QI Toolkit compared to other relevant QI Documents and Methodologies

The Toolkit states, "the World Bank Group and the National Metrology Institute of Germany (PTB) fully recognize the importance of QI as an ecosystem. The Toolkit is the first ever comprehensive QI diagnostics and reform guide".

Despite the still growing attention for QI, in actual practice, as well as conferences, (research) literature, and a great number of QI-related institutions around the world, there do not appear to be many parties engaged in operationalizing QI Methodology. Mainly, the number is limited to three significant players: the World Bank Group, UNIDO, and the German National Metrology Institute, the Physikalisch-Technische Bundesanstalt (PTB).

Apart from these parties, the Donor Committee for Enterprise Development⁴, the World Health Organization, and a variety of academics occasionally publish about QI Methodology, or closely related (QI-methodological) topics.

Ten Other Relevant QI Documents and Methodologies

The Toolkit is compared with the following other relevant QI documents and methodologies:

1. WTO-TBT Agreement, Guidelines, Principles, Decisions and Recommendations, Improving Governance for Regulatory Alignment (1995 - present)

WTO-TBT Agreement is an international treaty administered by the World Trade Organization. It was last renegotiated during the Uruguay Round of the General Agreement on Tariffs and Trade, with its present form entering into force with the establishment of the WTO at the beginning of 1995, binding on all WTO members.

The Technical Barriers to Trade agreement (TBT) was created to ensure the technical requirements (standards, technical regulations, and conformity assessments) do not lead to unnecessary obstacles to trade. All products, including industrial and agricultural products, are subject to the Agreement (except purchasing specifications and SPS measures, which are covered by other WTO agreements). The TBT Agreement recognizes that governments have the right to take the necessary measures for the protection of human, animal or plant life or health, but they should

⁴ The DCED is an independent forum where member agencies and others can come together to share experiences and to formulate a common understanding of good practice. This understanding is based particularly on empirical experience in the field: what has worked, and what has not. Member agencies are for example: IFC, ILO, FAO, UNIDO, OECD, and a number of national development agencies from countries such as Canada, UK, The Netherlands, Denmark and Sweden.



not discriminate against countries nor constitute disguised restrictions to trade. The TBT agreement includes provisions for the preparation, adoption and application of technical regulations and standards. Apart from the TBT, a separate agreement on the application of Sanitary and Phytosanitary Measures (SPS) is part of compliance to WTO signatories. SPS issues directly connect to a quality infrastructure of a country.

As a result the WTO-TBT Agreement has been further concretized by **Guidelines**, **Principles** (for example, of Openness, Transparency, etc.), continuously updated **Decisions and Recommendations**, and guidance on how to **improve regulatory alignment** with the WTO-TBT Agreement provisions, concerning standards, technical regulations, and conformity assessments.

The provisions concerning the preparation, adoption and application of standards are contained in Article 4 of the TBT Agreement and in the Code of Good Practice for the Preparation, Adoption and Application of Standards (the "Code of Good Practice"). In addition, Articles 2.4, 2.5, 5.4, and Paragraph F of Annex 3 of the Agreement promote the use of relevant international standards, guides and recommendations as a basis for standards, technical regulations and conformity assessment procedures. Articles 2.6, 5.5 and Paragraph G of Annex 3 emphasize the importance of Members' participation in international standardization activities related to products for which they have either adopted, or expect to adopt, technical regulations.

2. Trade Competitiveness Diagnostic Toolkit (Standards & Certification), International Trade Department of the World Bank (2011)

Implementation Toolkit, on Standards and Certification from a country's **trade competitiveness** perspective. While the policy prescriptions in any one country will be context dependent, certain broad **policy issues** need to be considered with regard to the standards environment. These are outlined below:

- Promoting awareness of standards
- Ensure that standards do not rule out local suppliers
- Role of lead firms in promoting standards
- Government may need to assist firms where lead firms do not
- Assisting small-scale producers
- Targeting low-income markets
- Harmonizing standards and developing countries participation in standards-setting bodies

3. Measuring the Impacts of Quality Infrastucture, National Metrology Institute of Germany, the Physikalisch-Technische Bundesanstalt (PTB) (2011)

The purpose of this report is to offer a better understanding about the kind of methodologies that researchers should employ when **assessing the impacts** of quality services. The key objectives to be achieved through the realization of impact assessment studies on quality infrastructure are:

- the identification of the range of positive and negative effects that they have in societies;
- to advise policy makers on how the accomplishment of their strategic goals may be facilitated through the development of quality services;
- the raise of awareness of consumers, firms and industries about the existence of such an institutional complex, i.e. how it shall be used as a platform where various economic agents are to participate and interact in order to address to a multitude of social, economic and environmental challenges;
- to learn ways of improving the functioning of the existing quality infrastructure.



Estimating the impacts of quality infrastructure services requires the adjustment of the general methodologies applied in impact studies to the particularities of quality systems. For the assessment of the impacts, the researcher has to bear the strong complementarities between services in mind, as they pose an obstacle when trying to isolate the impacts of individual quality services. The researcher has to question whether each impact is a consequence of the functioning of an individual service or from the integration of various services.

If the assessment study demonstrates that a quality service is unsuccessful to bring out the expected impacts, an enquiry is to be made to understand whether the complementarities between services were properly considered. This approach to impact studies on quality services states that they have to pass through three stages: establishment of an impact theory, verification of the theory, interpretation of the results. The impact theory defines the range of possible positive and negative impacts resulting from the functioning of a quality service or of a group of quality services, identifies the stakeholders and the impacts channels. To verify the impacts, a methodology has to be selected from a variety of possible approaches according to the information required, the information available, and the kind of impacts that are at stake. Interpreting the results requires situating the impacts in a context, in a strategy, listing the limitations of the study and proposing ways of optimizing the outcomes.

4. The National Quality Infrastructure (NQI) - A tool for competitiveness, trade and well-being, World Bank (2013)

This document **lists the components** of an NQI and reiterates the **importance of the role of governments** in developing an NQI. This includes sufficient **funding**, for which this document gives indications by estimating the costs and time involved in developing a NQI, itemized for the establishment of a National Metrology Institute, Legal Metrology practice, secondary calibration and testing laboratories, National Accreditation body, and a National Standards Body. Finally, it lists some **constraints** for the development and implementation of an NQI.

The NQI is the institutional framework that establishes and implements standardization, including conformity assessment services, metrology, and accreditation. While there are several approaches to the institutional set up an NQI, the best-practice approach is a decentralized system with the component organizations acting as legally autonomous units. Governments play a crucial role in designing, developing, and implementing an effective NQI. In countries with weak capacities, governments can lead coordination efforts by setting up technical committees, establishing testing facilities, and adopting standards into technical regulations for the purposes of consumer or environmental safety. They can help reduce transaction costs by gathering and disseminating information on standards and raising awareness about the benefits of adopting them. Government support has also proven critical in training technical personnel within NQI organizations.

There are several constraints to the successful development and implementation of NQI, including overly restrictive mandatory standards, conflicts of interest and political interference, lack of harmonization with international standards, limited financing, and lack of qualified personnel.

5. Supporting Business Environment Reforms, Practical guidance for Development Agencies, Annex: Supporting Quality Infrastructure in Developing and Transitional Economies, Donor Committee for Enterprise Development (DCED) (2014)

This annex focuses on how **donor and development agencies** can support QI in a broader context of **business environment reforms**. It is based on a series of deliberations by the DCED Business



Environment Working Group, including a detailed technical report. It presents common principles on the importance of a modern and market-oriented QI framework as an integral part of a sound business climate:

- Principle 1: Analyze demand and supply for quality services to set the right priorities
- Principle 2: Technical assistance to the national QI and regional QI systems must always be integrated in a holistic approach
- Principle 3: Embed technical assistance to building QI into a longer term planning, going beyond a single phase project
- Principle 4: Combine national and trans-national approaches
- Principle 5: Support integrated regional approaches
- Principle 6: Cooperation with organizations that are directly responsible for QI is most effective
- Principle 7: Strengthen demand and supply of quality services in parallel
- Principle 8: Promote institutional autonomy
- Principle 9: Consider maintenance and operating cost of expensive laboratory equipment
- Principle 10: Balance programme alignment needs with technical oversight and guidance
- Principle 11: Establish an active and diverse steering committee
- Principle 12: Strengthen financial and operational monitoring systems

6. Leveraging the Impact of Business Environment Reform: The Contribution of Quality Infrastructure - Lessons from Practice, Donor Committee for Enterprise Development (DCED) (2015)

Standardisation and technical regulation systems are integral to **business environment reform**. Although no definitive internationally accepted structure for the provision of QI services exists, good practices have evolved that should be taken into consideration. International recognition of QI organizations is vital for conformity assessment services to be accepted in the export markets. Sustainability of established QI organizations depends on two key issues: governments have a responsibility to provide funding for the fundamental QI institutions, i.e. metrology, standards and accreditation, and 'commercial' QI service delivery should be determined through appropriate market mechanisms, i.e. customers should have a choice as long as the technical competency of the service providers is ensured. Technical assistance to national QI and regional QI systems are not separate issues, but must always be integrated in a **holistic approach**.

Good practices in delivering QI related technical assistance should be considered within three main interrelated phases:

- 1. Project preparation and management
- 2. Strategic approaches to support quality infrastructure development
- 3. Support to quality institutions

7. Calidena Handbook 2.0., Physikalisch-Technische Bundesanstalt (PTB), GIZ (2016)

This Calidena Handbook consists of a **participatory methodology** developed and implemented by the Physikalisch-Technische Bundesanstalt (PTB) to stimulate quality in **value chains**; it aims to systematically and sustainably support the improvement of the NQI in developing countries and emerging economies. This handbook gives a structured overview of the approach and various key elements of the methodology. The **target audience** is project coordinators and members of steering committees of development projects, facilitators and representatives of quality



infrastructure organizations and of private sector organizations, who plan to work in the intersection between value chains and quality infrastructure. The methodology can be used in projects that aim to strengthen the **user orientation** of the NQI, and/or in value chain, initiatives that seek to address gaps in quality services.

8. Quality Policy, A Practical Tool, UNIDO (2018)

This practical tool consists of and is intended to be used in conjunction with, the UNIDO publications "Quality Policy Technical Guide", which details the various elements of a Quality Policy, and "Quality Policy Guiding Principles." These guides set out five core principles for Quality Policy development, namely Ownership, Inclusiveness, Coherence, Optimization and Sustainability. The tool adopts a process-based (step-by-step) approach, building on experiences obtained in different countries (each with their own specific context), and, in addition to answering the question of "*what*" to do, **guides the reader through the practicalities of the "***why***", "***who***" and "***how***" of each step**. It also examines some common or likely challenges and ways to overcome or mitigate them.

The Quality Policy development process is divided into **five key stages**, as follows:

- Stage 1: Do the groundwork Instill a sense of need for the Quality Policy, and promote ownership and coordination from within government
- Stage 2: Conduct strategic planning Define priorities and allocate resources
- Stage 3: Prepare the draft Quality Policy Provide for transparent consensus-building and assure coherence
- Stage 4: Lobbying, advocacy and approval Incorporate as part of national policy landscape
- Stage 5: Implement, monitor and review the Quality Policy Ensure effective and sustainable implementation

In the Annex, some **practical examples** of Quality Policy development are presented in a series of Case Studies.

9. Handbook for National Quality Policy and Strategy, World Health Organization (2018)

This is a handbook outlining the case for developing national **policy** and **strategy** on quality of health care, the process required to do so and supporting **tools**.

This handbook, and accompanying compendium of tools, provide direction on both the development process and content of national quality policies and strategies. It aims to facilitate development or refinement of these policies and strategies by policy-makers and practitioners who best know their unique country complexities. Users should see this document as a resource to help inform and structure quality policies and strategies responsive to the specific country needs while building on the guidance from existing literature, lessons from the field and expert consultation. The handbook is not a prescriptive guide, but rather a structured approach to help ensure development and implementation are comprehensive. To help users identify and access information, the handbook is divided into three color-coordinated parts. Part I focuses on quality policy development, Part II focuses on linked processes of strategy development, and Part III describes how to access and use supplementary tools to support the NQPS process.

The audience for this Handbook is diverse. The primary audience is those responsible for leading the development and implementation of national quality policies and strategies. A much wider set of stakeholders who actively participate in the national process will also benefit. Partners at the



national, regional and global level that are involved in providing support in quality improvement efforts will also stand to benefit from the content provided.

10. Measurement and performance of Quality Infrastructure - A proposal for a Global Quality Infrastructure Index, academic paper Buenos Aires & Duisburg (2019)

An academic discussion paper on the introduction of a Global QI Index that measures various aspects of the areas of metrology, accreditation, standardization and certification of products and services, both on the supply side (the international QI system and its services) and on the demand side (companies and other users of QI services). Information available on websites of international bodies as well as national and international organizations related to QI is collected. This measurement aims to establish country-level comparisons.

The data validate a positive correlation between the development of QI and the **economic competitiveness, performance and complexity** of a country. The data also show a **strong coherence** between the development of individual QI components. A country with high values in the field of metrology is usually also advanced in accreditation, standardization and certification. The opposite also applies.

How does the World Bank QI Toolkit compare with the best practices and opinions on QI Methodology?

Generally, the Toolkit has incorporated the latest insights of the WTO-TBT Guidelines, Principles Decisions and Recommendations, and guidance on how to improve regulatory alignment with the WTO-TBT Agreement provisions, concerning standards, technical regulations, and conformity assessments. The Toolkit does not however fully capture the multiple pathways approach articulated in the decision on the definition of an international standard from the WTO TBT Committee's Second Triennial Review. This weakness should be adjusted to accurately reflect the process of standards development rather than the position of the standard development body.

General Process Approach (Methodology 6 & 8)

The 2018 Quality Policy Tool of UNIDO approaches QI as a generic, 'holistic', process as does 'Leveraging the Impact of Business Environment Reform: The Contribution of Quality Infrastructure' from DCED. However generic and holistic, these methodologies both provide examples and case studies as further illustration. Compared to the Toolkit, which approaches QI more in a static, fragmented fashion in which practical examples or case studies are fully absent.

Limited, Value-Chain or Sector-wise, Process Approach (Methodology 7 & 9)

Simultaneously, some QI-methodologies focus on value chains, of specific sectors. This focuses on methodologies that provide guidelines or steps to develop QI in a value-chain, sector-specific approach, not only practically but also methodologically. This has consequences for the methodology, because it automatically means that QI is approached in a process-based manner, rather than a fragmented "building block" approach. In addition, it provides very practical knowledge 'from start to finish' for the stakeholder (farmer, trader, etc.) on how to comply with QI-matters in a process from produce to market⁵.

⁵ See for example "Quality Infrastructure supporting Nigeria to overcome the trade barriers of dried beans", https://www.youtube.com/watch?v=-L35Zz1F-YQ&t=85s.



A reason why a more focused approach may be effective is also that many developing countries only have a limited product-base for international trade, like beef from Botswana and bananas from the Caribbean. In such a situation, a value chain approach would probably fit better as a QI Tool than a full comprehensive QI assessment tool. In this context, it is also noteworthy perhaps, that the Commonwealth Standards Network funded by DfID, in Africa fairly recent has shifted their approach from assisting countries on comprehensive regulation-induced QI-reforms, towards a more practical, value chain approach focusing on specific sectors, such as fisheries, etc.

Outcome Approach (Methodology 2, 3 & 10)

Other methodologies included in this evaluation, concern with a more "impact assessment", or performance kind of approach of QI. To assess whether QI contributes to an objective of for example 'enhancing business environment', the effects of QI-interventions on business environment should be measured. Thus, it is more an outcome kind of approach⁶.

In addition, it is interesting to note that some QI-methodologies are expressly focused on achieving a specific objective, for example trade promotion, or - not surprisingly coming from the World Health Organization - improving public health. In fact, that is not uncommon, for example, many QI-projects around the world are conducted under the umbrella of a Trade Development Program.

'Single Stakeholder' Approach (Methodology 4 & 5)

Some QI-methodologies are expressly focused on a specific actor, or in other words a single stakeholder. QI Methodologies may focus solely on governments, highlight and specify their role in administration and funding responsibilities, as well as identify possible constraints and conflicts of interest, all from a government's perspective. Another QI Methodology solely takes the perspective of donor and development agencies.

Remaining issues

The Toolkit tends to use the EU understanding of many terms, the development of the Toolkit has happened in part under the auspices of the National Metrology Institute of Germany (PTB). The European use of terms deviates from the US model as well as the WTO-TBT Agreement (Decision of Second Triennial Review – definition of international standard). This EU-oriented introduction, however, does not seem to influence the outcomes or evaluative results of the Toolkit, as a limitation of the Toolkit remains expressing how to implement the terms described.

As for the validity of the Toolkit in the QI expert field: the Toolkit is perceived to be "the opinion" of author Martin Kellermann, which is why other parties in the field such as UNIDO and ISO have not endorsed it. These parties feel they have not been sufficiently consulted. This information was obtained from an established QI Specialist through an e-mail exchange of March 2020. This expert is

⁶ This is important, not in the least because QI for a large part is a commercial endeavor: standards can be bought, inspection fees must be paid, same for certification fees, even so storage, laboratory time, etc. Introducing (in fact, re-instating) Pre-export Verification of Conformity (PVoC) in Uganda via the Standards (Import Inspection & Clearance) Regulations, 2015 and the accompanying Administrative Directives on Implementation, 2015 (which introduces the application for an Import Clearance Certificate (PVoC Certificate), meant a rise in compliance cost for businesses of nearly 41 billion UGX annually. With the update of the Weights and Measures Rules in 2017 (update from the regulations of 2005/2007), and the Uganda National Bureau of Standards conducting nearly 800,000 verifications a year, compliance costs for business increased with nearly 8 billion UGX.



an ex-employee of the South African Bureau of Standards (SABS) who directly worked with Martin Kellermann (former executive director of SABS's non-commercial arm).

Conclusions

An Excellent Comprehensive Diagnostic, and Evaluative, Tool

The Toolkit is quite comprehensive in its approach. In this sense, it contains a lot of valuable information and is quite exhaustive, covering all aspects of the QI-domain. As such, it allows for an effective diagnostic, evaluation, or gap analysis for that matter, of a country's QI.

However, it is a Tool with an unclear Status...

The status of the document is not clear. The title labels it a 'Reform toolkit'. However, the document's foreword talks of a 'Guide'. Further, the document is referred to as a 'Diagnostic Tool' that uses a 'systematic Methodology'. In short, it is not clear what the writers have intended for the status of this document as a Guide, a Toolkit, a Diagnostic Tool, a Methodology or all of the above.

The status of the document has its consequences for its audience. Whether this document is a Guide, a Toolkit, a Diagnostic Tool, or a Methodology with the different characteristics and properties that go with each such forms, matters for how, and by whom this document is used.

....that is Prone to Subjectivity...

The Toolkit consists of many parameters and components. Many of which are highly dependent on what is, subjectively called 'expert knowledge'. The Toolkit and its proposed methodologies include no less than 10 parameters/components: Pillars, Building blocks, Classification Status, Implementation Status, Assessment of National Policy and Legal Environment with its own 3 parameters, Criteria (which are used as guidance in the expert assessment), Principles, and Level of Maturity. This may invite subjectivity into the use and application of the Toolkit, potentially yielding differing results when using this Toolkit in different circumstances/countries.

...while its usability is lacking attention for Process and Results

The Toolkit is excellent in providing an assessment of a static state of a country's QI. However, it does not pay much attention to the process of implementing QI, let alone the result of the assessment. With regard to the result, it is also important to what extent a country is going to conduct this assessment. This is relevant for the usefulness of the Toolkit, because, dependent on the intent or goal of a country, the Toolkit should be unambiguousness.

Special Note from the Field: Validity of the Toolkit in the QI Expert Field, according to an international QI Expert

The Toolkit document is essentially the opinion of Martin Kellermann, which is why others (UNIDO, ISO and others that were approached) have not endorsed it. They felt they were not adequately consulted.

Standards Alliance Needs Assessment Toolkit

NQI Assessment

#	Question			
	QUALITY ECOSYSTEM	SA Indicator	Scoring	Comments
	1 Does this country have a NQI Policy?	IR 1.1, IR 1.2	+0 No strategy +1 has active NQI +1 if regularly updated +3 has NQI Maximum of 5 points NQI Policy can be housed in multiple documents across multiple NQI actors, but should include a clear outline of roles and responsibilities	NQI Policy can be housed in multiple documents across multiple NQI actors, but should include a clear outline of roles and responsibilities
1	a When was it last updated	IR 1.2	Narrative, no score	
	2 Is the standards strategy in place and being implemented?	IR 1.1	Y/N (+1) +0 if no +2 for having a NSB	
	Is there a National Standards Body? Is the NSB internationally		+1 for participating internationally	
	3 recognized?	IR 1.1	Maximum of 3 points	
	4 Is there a National Metrology Institute? Is there at least one internationally recognized accreditation body active	IR 1.1	Y/N (+3)	
	5 in the country?	IR 1.1	Y/N (+1) Y/N (+2)	
	Is there at least one Conformity Assessment Body(s) active in the 6 country? Are the existing government quality institutes (Standards Body, Metrology Institute, CABs, Regulatory bodies) separate or does one entity play multiple roles?	IR 1.1	+1 if accredited Maximum of 3 points	
	This separation is key to identify conflict of interest issues in line with 7 international best practice.	IR 1.2	Y/N (+5)	
	8 Does the country have a national TBT/SPS committee? Does the country have active trade associations and consumer interest	IR 1.2, IR 3.1	Y/N (+2 for each enquiry point) Maximum of 4 points	
	9 groups?	IR 2.1	Y/N (+2 total)	list them and their effectiveness
	τοται		/21 (accreditation point is bonus)	

TRANSPARENCY	SA Indicator	Scoring	Comments
		Y/N (+1 each)	
1 Is there an active SPS/TBT Enquiry Point?	CCIR 1.1	Maximum of 2 points	
		+0 if no Enquiry point	
		+1 have notified in past 5 years	
		+1 notified this year	
	CCIR 1.1	+1 notify regularly (more than 12 per year)	
1a How active are the Enquiry Points?		+1 respond to enquiries within WTO window	
		+1 staff adequately trained	
		Maximum of 5 points	
		Narrative: Are Enquiry point staff using and capable of using the online enquiry	
		point resources?	

2 Where (in which ministry and office) are the Enquiry Points located?	CCIR 1.1	No score	
3 Is there a national policy or guidance on the notification process?	CCIR 1.1	Y/N (+1)	
4 Are draft regulations shared publically prior to finalization?	CCIR 1.1		
		+0 no national gazzette	
		+1 existing gazzette	
		+1 is it publically available (online?)	
4a Is there a national gazette? (does everyone call this a gazzette?)	CCIR 1.1	Maximimum of 2 point	How often is the gazzette updated and how often?
		+0 No online platform	
		+1 Has online platform that is functional/active	
Does the country maintain an online resource for sharing draft		+1 able to submit public comment online	
4b regulations for comment?	CCIR 1.1	Maximum of 2 point	How user-friendly is the online platform?
тот	AL	/12	
GENDER INCLUSIVITY	SA Indicator	Scoring	Comments
Are there mechanisms to support gender participation on Technical			
1 Committees (TCs)?	CCIR 2.1	Y/N (+1)	
			Use of the word all?
		+1 both genders are included in TCs	Is 25% fair? Is there a internationally-agreed upon
		+2 ICs have at least 25% participation from women	gender parity ratio we can use? (SDG goals?) Should this
2. Are both conders included in suisting TCs2	CCID 2.4	+1 At least 10% pt it's are chaired by women	be tiered?
2 Are both genders included in existing TCS?	CCIR 2.1	Maximum of 4 points	
		Total breakdown of TC chairs by gender	
		Gender breakdown of NSB (total and leadership roles)	
		Ratio of female to male labor force participation	
2a What is the current gender breakdown in the country?	CCIR 2.1	Ratio of female to male legislators, senior officials, and managers	
		+0 No gender policies	
		+1 Section in national standards stragety,	
		+1 national quality initiatives (UNI desease) ato2	
		+1 International gender Initiatives (ON decrees), etc.	
2 Are there existing policies that identify gonder inclusivity?		The manuary adopted of existing gender targets (ON SDGS)?	
5 Are there existing policies that identify gender inclusivity?	AI		
101		/ 10	

TECHNICAL REGULATIONS	SA Indicator	Scoring	Comments
1 Does country have existing regulatory framework?	IR 1.2	+3 policy exists +2 policy is followed Maximum of 5 points	
2 Is there a regulatory oversight body?	IR 1.2	 +0 No oversight +1 Centralized body for regulatory oversight Exists +1 Centralized body for regulatory review +1 has power to request regulations be reformulated based on RIA outcomes or that a RIA be reperformed +1 staff are capable of performing a RIA +1 actively training regulatory bodies on RIA Maximum of 5 points 	

3 Is there an existing mandate for regulators to perform RIA?	IR 1.2	Y/N (+1)	
Do regulators review the impacts of their regulations during the drafting	ng		
4 process?	IR 1.2	Y/N (+1)	
		+0 No to both questions	
r		+1 if private sector consulted	How many days are allotted for comment on draft
⁵ Is the private sector consulted in the development of regulations? Are		+1 if private sector can comment	regulations? Do they meet the WTO suggested
they able to review or comment on regulations before they are finalize	ed? IR 2.1	Maximum of 2 points	timeframe (60-90 days?)
5a			Add a description of opportunities when the private sector is consulted or invited to comment during the
How can the private sector be involved?	IR 2.1	Narrative, no score	regulatory development or approval process
If an arganization or government entity disagrees with a regulation arg			
6 there mechanisms to amond the regulation?		V/N (+2)	
Describe the process of creating a regulation (regulation development	to	1/18 (+2)	
7 finalization).		Narrative, no score	

/15

TOTAL

QUALITY INSTITUTES			
Conformity Assessment and Accreditation	SA Indicator	Scoring	Comments
		+0 for no existing strategy	
		+1 has certification strategy	
		+0.5 implementing certification strategy	
1 Does country have a conformity assessment or certification policy?	IR 1.2	Maximum of 1.5 points	
		+0 if No	
		+0.5 Yes	
		+0.5 if body is accredited	
		Maximum of 1 point	
2 Does the NSB or another body conduct certification of products?	IR 1.2	List bodies conducing this type of certification and what does each cover?	
		+0 No	
		+0.5 Yes	
		+0.5 if body accredited	
		Maximum of 1 point	
Does the NSB or another body conduct certification of			
3 systems/processes?	IR 1.1	List bodies conducting this type of certification and what does each cover?	
4 Is the national certification body regionally recognized?	IR 1.1	Y/N (+1)	
		+0 if No	
		+1 existing certification requirements for specific sectors	
		+1 information on how to obtain certification is available	
Is information on requisite national certification clear and easily		+1 certification requirements are clear	
5 available?	IR 1.1	Maximum of 3 points	
		+0 if No	
		+1 certifcation system exists	
Is there a national scheme in place for the certification of Quality		+0.5 plan to support SMEs in place	
6 Management Systems (QMS)?	IR 1.1	Maximum of 1.5 points	
Does the country participate in the ILAC Mutual Recognition Agreement			
7 (MRA) for certification?	IR 1.2, IR 3.1	Y/N (+1)	
			Qualitative measure for baselining, not included in this
8 Self score of effectiveness of certification system in country	IR 1.1	Quality institute self ranking on scale from 0 (does not exist) to 5 (world class)	SA2 scoring total
τοτο	1	/10	
IOTAL		/10	

Standards	SA Indicator	Scoring	Comments
1 Does this country have a national standards hody?	IR 1 1	+1 existing NSB +1 staff has capacity to execute daily functions +1 releases annual standards plan Maximum of 3 points	
	11 1.1	+0 for no strategy nor report +1 standards strategy or annual report exists +1 being implemented +1 regularly updated (at least every 5 years)	
2 Is there a national standards strategy or annual standards report?	IR 1.1	Maximum of 3 points +0 not independent +1 independent without oversight body +1 can adopt/revoke standards +1 can determine employee salaries +1 determines own funding	
Is the NSB independent (without oversight, can adopt/revoke standards,		+1 determines service fees	
determine worker salaries, determine own funding, initiate new actions,		+1 can intitiate new actions	
3 determine service fees)?	IR 1.2	SUM/6 for Maximum of 1 point	
4 Where is the NSB housed/located in the government?	IR 1.2	No score	
		+U NOT MEMBER OF ISU OF IEC +1 ISO membershin	
		+0.5 participates on at least one active ISO TC	
		+0.5 leads an ISO TC	
		+1 IEC Membership	
		+0.5 Participates on at least one active IEC IC	
5 Is the NSB a member to ISO/IEC and/or Regional Standards bodies?	IR 1.2. IR 3.1	Maximum score of 4	
Does the NSB have existing twinning relationships? Does the NSB recognize standards of regional/international standards	IR 2.1	No score, for baselining	
6 bodies?	IR 1.2, IR 3.1	Y/N (+3)	
		+1 NMI participates	
		+1 ACCreditation body	
		+1 foreign stakeholders	
		+1 academia	
		+1 widely advertised	
		+1 mechanisms to provide balanced participation	
NSB oversees robust standards development process (with openness.		+1 allows comments on draft standards	
7 balance, due process, consensus mechanisms)?	IR 1.2	SUM/9 for Maximum of 1 point	
		+0 not updated	
		+1 if reviewed	
8 How often are standards reviewed/updated?		+1 if reviewed on a regular basis (at least once per 5 years) Maximum of 2 points	
		+1 available in paper copy for purchase	
		+1 available in reading room	
		+1 available for purchase online	
9 Are standards available/accessible?		+1 available in online reading room	
			If yes, what percentage of the national standards
			catalogue mention use, adoption, application, reference, normative reference in whole or in part of international
10 Does the NSB have a formal policy to adopt international standards?		Y/N +1	standards
11 adopted standards)	IR 1.1	No score, for baselining indicator	
			Not included in this scoring, but a secondary metric for
12 Self-score of effectiveness of the NSB TOTA	L	Quality institute self ranking on scale from 0 (does not exist) to 5 (world class) /20	SAZ use

Metrology Institute	SA Indicator	Scoring	Comments				
		+1 if exists					
		+1 is accredited by international institute					
		+1 if recertify equipment every 3 years					
		+1 staff trained for testing					
		+1 staff trained for calibration					
		+1 staff trained for verification					
		+1 equipment regularly calibrated to working standards					
		+1/3 has implemented a formal QMS					
		+1/3 calibration labs accredited to ISO 17025					
		+1/3 participates in interlaboratory comparasions with other labs					
1 Does the country have national metrology institute (NMI)?	IR 1.2	Maximum of 8 points					
		+0 if no					
		+1 yes					
		+1 being implemented					
2 Does the NMI have a national metrology strategy?	IR 1.2	Maximum of 2 points					
		Quality institute self ranking on scale from 0 (does not exist) to 5 (world class)					
3 NMI self report score		Not included in this scoring, but a secondary metric for SA2 use					
	TOTAL	/10					
OTHER							
TBT/SPS at national/international level	SA Indicator	Scoring	Comments				
1 Which ministry/agency leads TBT issues?	IR 3.1						
2 Who is the lead organization to the WTO for TBT?	IR 3.1						
3 Does the country participate at the WTO TBT Committee?	IR 3.1						

KEY							
IR 1.1							
IR 1.2	Countries have knowledge about the value of using their national quality infrastructure Countries have enabling environment						
IR 2.1	Private Secor Participates in regulatory development						
IR 3.1	Increased awareness about TBTs in country						
CCIR 1.1	Countries NQI practices transparency						
CCIR 2.1	Gender incorporated into country NQI						

Sample completed questionnaire

#	Question	SA Indicator			Sco	oring			Points	Comments	
1	1a Does this country have a NOI Policy?	IR 1.1, IR	No strategy			Yes	Yes, & active	Yes, active, & reguarly updated		'NQI Policy can be housed in multiple documents across multiple NQI actors, but should include a clear outline of roles and responsibilities	
1	1.2 Loos this country have a NQI Policy?	1.2	0	1	2	3	4	5			
	1b When was it last updated?	IR 1.2			Narr	ative:					
2	2 Is the standards strategy being implemented?	IR 1.1	No	Yes							
			0	1	2	3	4	5			
3	Is there a National Standards Body?	IR 1.1	No			Yes					
			0	1	2	3	4	5			
4	Is there a National Metrology Institute?	IR 1 1	No			Yes					
			0	1	2	3	4	5			
F	la thora a National Acaraditation Dady2	IR 1.1	No	Yes							
5	5 Is there a National Accreditation Body?		0	1	2	3	4	5			
	6 Is there a National Conformity Assessment Body(s)?	dy(s)? IR 1.1	No		Yes						
6			0	1	2	3	4	5			
	Are the existing government QI bodies		No					Yes			
7	7 7a separate or does one entity play multiple roles?	IR 1.2	0	1	2	3	4	5			
	If one QI body plays multiple roles, are		No or N/A			Yes		-			
	7b (Standards, Conformity Assessment,	IR 1.2	0	1	2	3	4	5			
	8 Does the country have a national TBT/SPS committee?	10.4.0.10			1 enquiry point		2 enquiry points				
8		ir 1.2, ir 3.1	0	1	tor IBI/SPS	3	4	5			
9 Does the co and		IR 2.1	No		Yes						
	Does the country have active trade associations and consumer interest groups?		0	1	2	2	4	5			
Total			0	Ŧ	2	5	7	3	0	/21 (accreditation point is bonus)	

0 /21 (accreditation point is bonus)