

MEASURING YOUTH COMPETENCIES ACROSS CONTEXTS: LESSONS FROM IMPLEMENTERS ON HOW TO ADAPT SOFT SKILLS MEASUREMENT TOOLS

YOUTHPOWER CROSS-SECTORAL SKILLS FOR
YOUTH (CSSY) COMMUNITY OF PRACTICE
TECHNICAL BRIEF



Photo by Cindi Fabiola Mejia

INTRODUCTION

Measuring soft skills is a task that continues to challenge program implementers and monitoring and evaluation (M&E) experts around the world. Growing evidence highlights the importance of these skills in contributing to youth development, but their intangible nature makes them difficult to assess. This challenge is compounded when implementing soft skills programs across geographic and cultural contexts, in which cultural differences can play a role in the fundamental definitions and manifestations of these competencies.

The purpose of this technical brief is to gather and synthesize lessons learned from implementers' experiences adapting soft skills measurement tools in diverse geographic and cultural contexts through a brief scan of the literature and three implementer case studies. It is not intended as technical guidance for tool adaptation, but instead as a means of leveraging the knowledge and experiences of community of practice (CoP) members to improve practice by identifying common challenges and lessons learned.

The brief consists of four sections. First, the importance of soft skills to youth development is outlined along with the broader challenges associated with measuring

such skills. Second, challenges specific to measuring soft skills across contexts are sourced from relevant literature. Third, the brief presents three case studies from program implementers working to measure soft skills across contexts. Finally, common lessons learned on measurement tool adaptation are discussed.

THE IMPORTANCE AND MEASUREMENT OF SOFT SKILLS

Soft skills¹ refer to a broad set of skills, behaviors, and personal qualities that enable people to effectively navigate their environment, work collaboratively with others, perform well, and achieve their goals. These skills are applicable across sectors and complement the acquisition of technical, vocational, and academic skills (Gates et al., 2016). A growing body of research indicates that soft skills are important to school and workplace success as well as to reductions in risky behavior (Wilson-Ahlstrom et al., 2014). Evidence shows that when skills-based programs achieve impacts in these areas, they also make progress on more traditional academic measures such as grades and test scores (Wilson-Ahlstrom et al., 2014).



Certain soft skills have been identified as key to specific outcome areas. For example, a recent study by Gates et al. (2016) found that positive self-concept, self-control, and higher-order thinking skills effectively support youth advances across the workforce success, violence prevention, and sexual and reproductive health outcome areas. Additionally, social skills and communication emerged as skills with strong cross-sectoral support, according to the study. Moreover, empathy and goal orientation were identified as particularly important skills specific to violence prevention and sexual and reproductive health outcomes, respectively.

Despite ample evidence on the benefits of soft skills for youth development, challenges persist in measuring these skills. A recent review and inventory of soft skills measurement tools by YouthPower Action (Galloway et al. 2017) listed the following methodological challenges in the field of soft skills measurement:

- Balancing technical considerations such as reliability, validity, and measurement invariance²
- Using tools to reliably measure change in skills over time, when measuring a soft skill at a single point in time is itself challenging
- The prevalence of self-report methods that are known to suffer from biases
- Lack of implementer inclusion in tool design
- *Developing or adapting tools for use across cultures and contexts with limited resources*

This technical brief focuses on the last challenge, synthesizing lessons learned from implementers' experiences adapting measurement tools in different contexts.

STEPS TO EFFECTIVE PROGRAMMING ON YOUTH, PEACE, SECURITY AND PVE/CVE

Several challenges in adapting soft skills measurement tools across geographic and cultural contexts have been captured in youth development literature. These include self-reporting bias, a lack of common terminology and understanding regarding specific soft skills across contexts, and relative value of particular soft skills across cultures.

¹The nomenclature around these skills is still highly debated, with terms like "life skills" and "cross-sectoral skills" also used interchangeably with "soft skills." For this brief, we have decided to use the "soft skills" framing to stay consistent with the literature that is referenced as well as the vocabulary generally used by consulted stakeholders.

SELF-REPORTING BIAS

Evidence from a scan by Olenik et al. (2013) and a review of soft skills measurement tools indicates that most assessments rely heavily on self-report measures, a commonly used but flawed approach to capturing soft skills levels. A similar scan of soft skills measurement tools conducted by YouthPower Action researchers found that even the 10 highest-rated tools studied relied heavily on self-report measures (Galloway et al. 2017)³. Per the reviewers, the accuracy of self-report data can be compromised because youth often do not self-report negatively viewed perceptions or behaviors. Additionally, these measures also suffer from reference and social desirability biases and are limited in their ability to track individual changes over time.

LACK OF COMMON TERMINOLOGY AND UNDERSTANDING

As recently noted by Galloway et al. (2017), most soft skills measures have been built and tested in developed countries. As a result, these assessments are inherently biased towards western cultural terminology and understanding of soft skills concepts, which may not align directly with those of developing country contexts. Complex concepts such as "critical thinking," "conscientiousness," and "self-motivation" may not have the same meaning to people of other cultures (Kautz et al., 2014). This is particularly important given the focus of many assessments on self-reported data: If youth taking soft skills assessments do not have the same reference point about each concept, they will not rate themselves accurately or consistently. The distortion brought on by a lack of common expression around concepts across cultures—known as cultural reference bias—also hampers the ability of program implementers and evaluators to compare results across programs and cultural contexts.

DIFFERENCES IN THE RELATIVE VALUE OF SOFT SKILLS ACROSS CULTURES

Differences in the importance ascribed to specific soft skills may also have implications for how those skills are assessed in different contexts. For example, research conducted on key education stakeholders' attitudes and perceptions of soft skills in Mexico, South Africa, Kenya, and the Philippines show that there are clear differences between countries in the types of soft skills that are highly valued and emphasized (Care et

² Measurement invariance is the statistical property of a measurement that indicates that the same underlying construct is being measured across groups or across time.

³Per Galloway et al. (2017) reference bias occurs when "frames of reference differ by individual according to their social group norms" while social desirability bias occurs when "individuals provide answers that they perceive to be 'desirable' but are not accurate."

al., 2017). While education stakeholders in South Africa identified communication and listening as key soft skills for learners, stakeholders in Mexico pointed to skills related to the self, including self-awareness and self-regulation, as the most important soft skills (Care et al., 2017). These differences could impact how youth perceive and respond to different types of soft skills. Furthermore, due to differing perspectives of soft skills across contexts, some items may be considered 'sensitive,' making them more difficult to assess (Olenik et al., 2013).

MEASUREMENT TOOL ADAPTATION CASE STUDIES

Despite the numerous challenges associated with measuring soft skills across contexts, program implementers and psychometric experts have developed innovative approaches to ensure that measurement tools are relevant to specific cultures and settings.

This brief reveals lessons from such efforts in consultation with members of the CSSY CoP involved

in soft skills measurement across contexts. CoP members were asked to provide descriptions of their respective assessment adaptation and development processes. Based on the input and data collected, three implementers were selected for case studies that summarize their efforts: Education Development Center (EDC), World Vision, and the International Youth Foundation (IYF). At least one staff member from each organization was interviewed and some also provided supporting documents to inform their respective case studies.

While not comprehensive, the case studies share each organization's rationale and process for tool adaptation. Each organization approached this challenge in a different way. EDC enhanced a pre-existing measurement tool with questions that strive to pre-empt cultural bias, World Vision developed a contextualization process that consults youth early on in development and can be applied to more than one measurement tool, and IYF developed a brand new digital and adaptable assessment tool.

1. Education Development Center (EDC): Grounding assessment data through anchoring vignettes to offset cultural biases in skills assessment

EDC's youth development initiatives focus largely on skill development and workforce readiness. Its *Work Ready Now!* curriculum has reached over 100,000 youth in more than 20 countries, and has even been adopted as the national work readiness curriculum in Rwanda and Macedonia. Like other program implementers working in this space, EDC has faced challenges measuring soft skills across its youth employability programs. Soft skills are not only inherently difficult to capture through measurement, but the broad geographic range of EDC's programs means that cultural differences can play a major role in how youth interpret or assess their own capabilities. Confronted with the need to quantify the impact of its programming across contexts, EDC strengthened a pre-existing soft skill measurement tool with new questions that made it more adaptable. In 2016, together with then-partners Professional Examination Services (ProExam) and The Akilah Institute for Women, and through the USAID-funded Workforce Connections initiative led by FHI360, the group developed a new soft skills tool called the *Anchored BFI*.

Anchored BFI is a psychometrically sound soft skills assessment meant for multiple developing-country contexts. The tool is based on a widely accepted framework for soft-skills development called the 'Big Five' model. It builds upon the Big Five Inventory-44, a self-report instrument developed to measure an individual's competencies in the Big Five core skills: conscientiousness, agreeableness, neuroticism, openness, and extraversion.

EDC and partners ensure the Anchored BFI is adaptable across contexts by consulting youth during the measurement tool translation process and including items which limit the "fakeability" inherent in self-reporting and cultural reference bias. The following steps were taken by EDC in development and adaptation:

- Cognitive interviews with local youth: As the tool was being translated into a new language, EDC's local assessors (who speak the same language as local youth) went through the tool with at least two youth in each country (one male and one female) asking them to react out loud as they read each question to identify items that did not make sense or may not have been directly relevant to them as stated. In some cases, EDC probed youth on hard-to-translate concepts, such as "conscientiousness." Revisions based on youth feedback were then made to the assessments' text.

- Situational judgement tests (SJTs) to limit “fakeability:” Measurements that are based heavily on self-report data are unreliable, as respondents can often guess a favored response and inflate their scores. To limit “fakeability,” EDC included SJTs—questions which present users with a hypothetical scenario in which the “correct” response is not always obvious—in the Anchored BFI. These questions allow for a more accurate measurement of respondents’ competencies across a number of skills.
- Anchoring vignettes to correct for cultural reference bias: An anchoring vignette is a type of survey question that corrects for cultural and contextual bias by gauging an individual’s understanding of how a soft skill is ideally expressed, and then adjusting their self-reported ratings. For example, given that the term “conscientiousness” may have different cultural meanings across contexts, an anchoring vignette would give the respondent three hypothetical scenarios related to that skill set and ask them to rank these scenarios based on what they believe to be the best expression of being “conscientious.” Based on these responses, the respondent’s answers can be rescaled to account for their own understanding of what excelling at that skill means.

EDC’s Anchored BFI has so far been implemented within the organization’s youth programming in Rwanda, Honduras, and the Philippines. Going forward, EDC hopes to use the tool for pre-and post-program evaluation data collection.

2. World Vision: Starting survey development from scratch with local youth experts

World Vision’s wide-reaching youth development work, based on the Positive Youth Development (PYD) framework, consists of an integrated approach to strengthen economic and livelihood empowerment, citizenship-building, child protection, education, and health outcomes. The organization uses the Search Institute’s Developmental Assets Profile (DAP)⁴ tool to assess youth wellbeing. This tool was generated in the U.S. and looks at adolescents holistically, assessing both their internal assets (e.g., positive values, social competencies, identity), and external assets (e.g., boundaries and expectations, support, empowerment). While effective in the U.S., World Vision found the need to create culturally appropriate, local-language versions of the tool to implement across different contexts. Attempts at translating and adapting the DAP led World Vision to develop and formalize its own contextualization process. This process has since been applied to the organization’s soft skills programming in order to adapt measurement tools which assess soft skills such as critical thinking, communication, and conflict resolution.

The process relies heavily on input from local staff and youth, and it includes the following steps for adaptation into each new language, country, and/or region:

- Qualitative research on skills concepts: World Vision’s contextualization team, which consists of measurement experts, translators, and youth practitioners, conducts focus group discussions with local youth to brainstorm around soft skills concepts such as “critical thinking.” The session is loosely facilitated and allows youth to express themselves freely, create mind-maps, and explore what it means to exemplify the concept at hand. This helps World Vision in two ways: in outlining the cultural relevance of the targeted concept for local youth, and in highlighting useful local language and terminology around that concept.
- Translation exercise with child experts: A group of 12 “youth experts” (local youth who speak both the local language and English) translate survey items. The idea is not to translate items verbatim, but instead have youth contribute ideas for how to state each item in a way that retains its original meaning to local youth.
- Staff expert translation screening: The contextualization team reviews youth experts’ translated versions of the tool and assesses each item based on three criteria: Does the translation stay true to the items’ original meaning? Is the translated text the easiest way to express this item for all target ages? And will the language used make sense for respondents across the country? Items are revised, back-translated into English as a final quality check, and finalized based on the team’s agreement.
- Pilot beta and field testing: The survey is then incorporated into a beta version of the measurement tool and piloted with 100 youth, including those from more marginalized settings, to ensure relevance across demographic lines. Guided by the pilot test data, adjustments are made before the measurement tool is field tested with an additional 300 youth (including test-retest), readjusted, and finalized.

⁴Scales et al. (2016) provide a peer-reviewed analysis of DAP’s implementation around the world, including lessons learned from these efforts.

Over the last seven years, World Vision has refined this contextualization process and created local language versions of the DAP in 15 countries. More recently, the process was utilized to create soft skills measurement tools in Cambodia and Honduras, which will be used primarily for program evaluation.

3. International Youth Foundation: Testing for workforce readiness across contexts by mitigating “fakeability” and piloting across six countries

The International Youth Foundation (IYF) has worked to improve youth livelihoods through the development of leadership, technical and life skills for over 25 years. Its flagship youth program, Passport to Success (PTS), works with youth ages 15 to 29 in a diverse range of settings, including secondary schools, vocational technical schools, and non-profit organizations. The program’s curriculum focuses on soft skills such as self-confidence, emotional management and conflict resolution, career planning, respect, cooperation and teamwork, project planning, and other competencies critical in both school and at work. It has been implemented in more than 50 countries.

While IYF programs have had great reach, capturing their transformative impact on youth work readiness through soft skills assessments is an ongoing challenge. After attempting numerous approaches and measurement tools, IYF recently decided to develop a rigorous psychometric assessment tool for measuring soft skills work readiness.

Beginning in 2015, IYF partnered with ProExam to develop this tool per international testing standards. The assessment focuses on six core soft skill constructs (also derived from the ‘Big Five’ model) consistently identified as critical by employers and industry associations and reflected in multiple research-based skills competency models. These are: conscientiousness, integrity, cooperation/teamwork, emotional responsiveness, open-mindedness/creative orientation, and social/leadership orientation. Seeking to address concerns about “fakeability” and cultural-reference bias that often accompany self-reported surveys, the new assessment includes multiple item types, such as forced-choice and situational judgement, as well as Likert-style self-report items.

From December 2016 to March of 2017, IYF and ProExam piloted the assessment with 1,640 youth in six countries: Mexico, El Salvador, Indonesia, India, South Africa, and Jordan. The pilot data is key to determining the cross-cultural applicability of the assessment, particularly across different demographics variables. The following steps were taken in developing and piloting the new tool:

- Item testing and translation review pre-pilot: Youth participants reviewed an initial batch of items as the tool was created to inform the test developers on crucial issues such as reading level and terminology. Additionally, the translations for Spanish and Bahasa were reviewed by native speakers to ensure accuracy. | Assessment comparability evaluation: Data from the six country pilots was evaluated for comparability by country and by language. Items that did not perform consistently across cultures were identified and will be removed from the final version of the measurement tool.
- Post-test user focus groups: These were conducted to attain qualitative information on the user experience and to confirm the cultural appropriateness of the translations among pilot participants.
- Local validity assessment: With the support of in-country subject matter experts, IYF conducted a review of individual items and the assessment form focusing on cultural and linguistic considerations.
- The measurement tool, which is fully digital and implemented online, has been designed for multiple uses. It can be implemented:
 - As a diagnostic tool, for example, by educators to identify skills in need of remediation or further development
 - As a summative tool (or capstone) to evaluate the impact of skills development programs (such as Passport to Success)
 - As a formative tool to provide feedback to youth and educators on their progress through development programs, and to help youth communicate their skills to prospective employers

IYF is currently working with ProExam (now part of ACT Inc.) to augment the data collected in the pilot by continued field testing with approximately 10,000 more young people. This process will provide additional evidence of the validity and cross-cultural applicability of the assessment, which will be available globally in early 2018.

COMMON LESSONS LEARNED

Based on the case studies above, the CSSY CoP has identified five common lessons learned that we believe are relevant to others currently adapting soft skills measurement tools across contexts.

1. A diverse, collaborative team is crucial to adapting tools across contexts

The adaption of skill measurement tools requires a mix of technical expertise, local know-how, critical thinking, and adaptability. Each of the three program implementers highlighted in their interviews the importance of having a diverse team working together to ensure that measurement tools are effective in different contexts. As noted by EDC, a collaborative working environment in which team members bring their own strengths to the table is crucial. EDC, World Vision, and IYF each rely on teams that are both technically and culturally diverse and include technical psychometricians, translators, and program implementers. Local staff and program implementers, who are intimately connected to youth on the ground, were noted as particularly crucial to ensuring assessments are contextually relevant.

2. Involving youth early on and throughout the adaptation process can help ensure local relevance

To different extents, all three organizations involved youth early in the development of their measurement tool. World Vision's contextualizing process requires engagement with youth to brainstorm about skills concepts before survey items are developed. EDC and IYF's cognitive interviews with youth ensure that translated survey items are relevant and demonstrate an early commitment to youth engagement. The value of young people's contribution to ensuring assessment tool validity was observed first-hand by our interviewees. A colleague from World Vision noted that in Honduras, for example, the team's professional translator was amazed that the survey items composed by youth experts were more direct and relevant than those he had originally drafted.

Moreover, better tools were not seen as the only positive outcome of youth engagement. Interviewees noted the enthusiasm youth had for participating throughout the adaptation process, including through focus groups, translation exercises, cognitive testing, and piloting. As noted by an IYF colleague, engagement with local youth also allows the youth to feel they have a voice in the questions that are being asked. This level of empowerment is one that IYF hopes to translate to youth's post-assessment engagement as well, by making

its work readiness tool one that youth can leverage to advocate for themselves with potential employers.

3. The process and practicality of adapting measurement tools is also varied across contexts

The most noticeable challenges related to measurement tool adaptation across contexts are associated with language translation and cultural relative bias. However, interviewees also noted challenges that went beyond terminology and instead involved the process of respondents completing assessments. For example, in testing some of their anchoring vignettes, World Vision found that the concept of an arbitrarily named hypothetical person or "friend" whose skills competencies were to be ranked was one that was not culturally intuitive to youth in Cambodia. Similarly, EDC found that secondary school youth in Rwanda (whose literacy levels should have enabled them to complete the Anchored BFI assessment with little difficulty) took an extremely long time to finish the assessment, compromising its accuracy. These examples demonstrate that adaptation issues also extend to the actual process of test-taking: How a measurement tool is worded, organized, and proctored should also be considered as it is adapted to fit a new context. In all three case studies, program implementers kept a close eye on contextual challenges throughout piloting. Some data collection for tool development feedback, such as IYF's user experience focus groups, was done to identify issues related to tool adaptation.

4. Adaptation is sometimes needed within countries to account for local and demographic nuances

In some instances, survey tools need not only to be adapted across countries, but also for specific regions within in a country, in order to truly account for the diversity of youth in that context. Though each of the three implementers highlighted their steps for adapting tools across countries, they also paid special attention to demographic and gender differences in respondents' experience with assessments. Translation checks and pilot testing for all three of the case studies cited in this brief were careful to consider youth's background, income level, and gender status in ensuring each tool's relevance. EDC noted, for example, that the tool they developed for youth in Manila had to be tweaked when applied to rural youth from Mindanao, in the Philippines, to account for different circumstances. Interviewed implementers noted that local staff and program implementers are crucial to identifying important nuances within countries.

5. Considering the purpose of an assessment is important when deciding on adaptation activities

Keeping assessment tools' purpose in mind is crucial to ensuring skills measurement is possible given the limited time and resources available to program implementers. The difference in each assessment's objectives, resources and time spent in adaptation can vary. For example, World Vision's contextualization process can take three to six months, while EDC's Anchored BFI and IYF's work readiness assessment have each taken years to develop and pilot. As they discussed their individual approaches, each of our case study interviewees also

made sure to note their limitations. EDC's grounding on a pre-established 'Big Five'-based assessment, World Vision's limitations around the use of this process for its own programs, and IYF's focus on work readiness and digital implementation all allowed these programs to define their scope and keep their objectives in mind when developing their assessment tools. In doing so, they could tailor adaptation activities to their purpose and prioritize their limited resources.

REFERENCES

- Care, E., Kim, H., Anderson, K., and Gustaffson-Wright, E. (2017). *Skills for a changing world: National perspectives and the global movement*, The Brookings Institution. <https://www.youtheconomicopportunities.org/resource/8531/skills-changing-world-national-perspectives-and-global-movement>
- Galloway, T, Lippman, L., Burke, H., Diener, O., and Gates, S. (2017). *Measuring Soft Skills & Life Skills in International Youth Development Programs: A Review and Inventory of Tools*. Washington, DC: USAID's YouthPower Implementation IDIQ-Task Order 1, YouthPower Action. <http://www.youthpower.org/resources/measuring-soft-life-skills-international-youth-development-programs-review-and-inventory>
- Gates, S., Lippman, L., Shadowen, N., Burke, H., Diener, O., and Malkin, M. (2016). *Key Soft Skills for Cross-Sectoral Youth Outcomes*. Washington, DC: USAID's YouthPower: Implementation, YouthPower Action. <http://www.youthpower.org/resources/key-soft-skills-cross-sectoral-youth-outcomes>
- Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., and Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success (No. w20749)*. National Bureau of Economic Research. http://forumfyi.org/files/soft_skills_hard_data_single.pdf
- Olenik, C., Zdrojewski, N., Bhattacharya, S. (2013) *Scan and Review of Youth Development Measurement Tools*. USAID. Prepared by JBS International. <https://www.usaid.gov/sites/default/files/documents/2155/USAID%20Life%20Skills%20Measurement%20Review%20FINAL%20EXTERNAL%20REPORT.pdf>
- Scales, P.C., Roehlkepartain, E.C., & Shramko, M. (2016). *Aligning Youth Development Theory, Measurement, and Practice across Cultures and Contexts: Lessons from Use of the Developmental Assets Profile*. Child Indicators Research. DOI 10.1007/s12187-016-9395-x. <https://link.springer.com/article/10.1007/s12187-016-9395-x>
- USAID (2016). *Measuring Youth's Soft Skills Across Cultures: Evidence from the Philippines & Rwanda*. Produced for Workforce Connections by Education Development Center (EDC), Professional Examination Services (ProExam), and Akilah Institute for Women. <http://idd.edc.org/sites/idd.edc.org/files/Anchored%20BFI%2010%20page%20report.pdf>
- Wilson-Ahlstrom, A., Yohalem, N., DuBois, D. L., Peter, J., and Hillaker, B. (2014). *From soft skills to hard data: Measuring youth program outcomes*. Forum for Youth Investment. http://forumfyi.org/files/soft_skills_hard_data_single.pdf

This brief was developed by Daniel Plaut, Caitlin Moss, and Shubha Jayaram (R4D) in consultation with members of the Cross Sectoral Skills for Youth (CSSY) Community of Practice. Case studies were developed through interviews with program implementers at EDC, IYF, and World Vision. Special thanks go to Rebecca Povec Pagel (EDC); Teresa Wallace (World Vision); and Lara Henneman, Sheerin Vesin, and Karen Phillips (IYF). For more information on YouthPower or the CSSY CoP, please visit <http://www.Youthpower.org> or email the project at info@youthpower.org

CONTACT

YouthPower Learning

Making Cents International
1350 Connecticut Ave, NW
Suite 410

Washington, DC 20036 USA
www.YouthPower.org

 @YPLearning

 YouthPowerLearning

USAID YouthPower Learning generates and disseminates knowledge about the implementation and impact of positive youth development (PYD) and cross-sectoral approaches in international development. The project leads research, evaluations, and events designed to build the evidence base related to PYD. Concurrently, YouthPower Learning employs expertise in learning and knowledge sharing to promote engagement and inform the global community about how to successfully help transition young people into productive, healthy adults. YouthPower Learning supports the implementation of the 2012 USAID Youth in Development Policy to improve capacity and enable the aspirations of youth so that they can contribute to, and benefit from, more stable, democratic, and prosperous communities.

This brief is made possible by the support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of Making Cents International through YouthPower: Evidence and Evaluation Task Order 1 (YouthPower Learning) AID Contract # AID-OAA-I-15-00034/AID-OAA-TO-15-00011. This report does not necessarily reflect the views of USAID or the United States Government.