Lee Kuan Yew School of Public Policy
National University of Singapore

Policy Analysis Exercise, Master in Public Policy, 2017

April 14, 2017

Citizen-driven measurement of the Sustainable Development Goals

How can Perceptions Data measurement contribute to the 2030 Agenda for Sustainable Development?

Anita Dhillon

Ian Reinert

Daniel Willett

Executive Summary

The UNICEF Global Innovation Centre launched an SMS-based platform called U-Report in 2011 to connect young people with their governments and provide information on topics important to them. Currently U-Report has 3 million users (and is rapidly growing) making it an interesting candidate for enhancing the reporting on the Sustainable Development Goals in an open, inclusive, transparent and participatory way. It could effectively achieve this through the use of Perceptions Data (defined in this paper as expressed views of people on a particular topic) when reporting progress towards the Sustainable Development Goals (SDGs). Therefore, this paper, undertaken in partnership with UNICEF, attempts to fill an urgent gap in understanding how to effectively measure people's perceptions to enhance the reporting of progress towards the Sustainable Development Goals (SDGs), and especially how these lessons could be applied to U-Report.

Perceptions Data gathering and utilization has the ability to enhance citizen participation and understand the underlying causes and mindsets that lead to a lack of progress towards certain social indicators, like those targeting gender norms. Currently, Perceptions Data is being collected across traditional face-to-face surveys and newer, technology-based platform utilizing SMS and instant messaging. One key issue being faced is that prior to this new technology enabling systematic collection of Perceptions

Data, the existing available findings weren't really on the radar of key decision makers and were thus not something they were held accountable to by the wider public (although this is changing with the recent political upheavals in the Western world). Additionally, with these new methods of data collection, it is important to balance their potential cost effectiveness with the quality and robustness of data generated.

Given this context, the paper fills a gap in existing research through identifying key existing tools that measure perceptions and evaluates their effectiveness in doing so by using an Analytical Framework that draws not only from a theoretical and research-based perspective, but also includes a focus on practical applicability of results. Specifically, the evaluation targets four key standards to determine effectiveness: Representation, Timeliness, Measurement, and Credibility which is then used to identify key strengths of tools and areas that can be improved upon to better measure perceptions to determine progress towards the SDGs.

From the analysis it is concluded that there exist two typologies of tools showcasing the trade-off between the Representation and Timeliness standards and that while the tools designed for representation are more direct and controlled in their measurement of perceptions, biases still remained across both typologies. Additionally, information on the questions and tool methodologies are generally open with certain tools more comprehensive in their coverage of SDG indicators. These findings lead to three key recommendations being made to answer the research question:

- **Confront the trade-off between timeliness and representation:** This includes greater use of mixed methodology approaches combining existing tools with new techniques to examine for biases, working with partners to build and test representative sample frames, and utilising clever design to address literacy, non-responses, and attrition within the sampling frame;
- **Use a framework for effective question design:** In building an Analytical Framework capable of assessing what makes an effective tool, a significant number of questions were analysed, providing clear examples and a framework for how to gather perceptions on youth related targets and indicators. This framework is designed to help practitioners confront framing and recall effects, social desirability and recall biases, and other potential pitfalls of question development and design; and
- **Make SDG perceptions data more accessible:** These tools are unlikely to replace existing measurement strategies and techniques, and nor should they, but they have a number of strengths that can contribute to the existing measurement ecosystem. This involves bringing data collected into a single, easy to access open data portal, and examining how different data sets can be effectively combined to build deeper understanding of the complex constructs that are the focus of the Agenda for Sustainable Development.

Contents

I. Problem Definition and Context

I.a. Background, Context and Research Question

The establishment of the Millennium Development Goals (MDGs) represented a watershed moment in global development efforts, with an articulate global agenda and agreed metrics allowing for significantly greater international coordination. Criticism, however, arose that these MDGs were not responsive to local realities (Vandemoortele, J, 2011). Following this, the development of the 2030 Agenda for Sustainable Development has adopted a more participatory and inclusive approach to both defining the problem and reviewing national progress towards it (UN General Assembly, 2015). A crucial way in which it aims to do so is by promoting the use of Perceptions Data (defined in this paper as expressed views of people on a particular topic) when reporting progress towards the Sustainable Development Goals (SDGs).

The Agenda sets a roadmap for the measurement and reporting of progress towards the SDGs and its indicators. Along with a focus on rigour, evidence-based approaches and data disaggregation, there is a strong push for holistic measurement of progress, as highlighted by the following excerpt:

> A robust, voluntary, effective, participatory, transparent and integrated follow-up and review framework will make a vital contribution to implementation and will help countries to maximise and track progress in implementing this Agenda in order to ensure that no one is left behind. (UN General Assembly, 2015).

As part of a broader push to engage local people in development, UNICEF launched U-Report first in Uganda in May 2011. U-Report began as an SMS-based platform for youth to shape their futures by providing thoughts and opinions on issues and topics important to them. It has since been scaled up to 34 nations. The U-Report platform aims to use this information to connect young people with their representatives, improve delivery of social service programmes and draw the attention of national governments and the UN to urgent issues. For example U-Report is a leading provider of sexual health information in Mozambique, enabled partially by the level of confidentiality afforded by the use of technology ("How SMS," 2016). With 3 million 'U-Reporters' on this platform, there is interest in examining whether and how it can be used to measure and report on key sustainable development indicators in an 'open, inclusive, participatory and transparent' way. In this context and working in partnership with UNICEF, the research questions being analysed in this paper are:

- **What makes a tool effective and which current tools have been the most effective, in measuring people's perceptions?**; and
- **How can these tools for measuring perceptions - especially U-Report - be used to enhance the measuring and reporting of SDG progress?**

To focus the analysis of perceptions measurement for the SDGs, only a subset of its 230 indicators were considered for this paper. The responsibility for reporting on these indicators lies with national governments, however, the United Nations Statistical Commission formed the Inter-agency and Expert Group (IAEG) in March 2015 to develop an indicator framework for the goals and targets established as part of the Sustainable

Development Agenda at the global level, and support its implementation ("IAEG-SDGs," n.d.). As the leading UN agency for children, UNICEF has a key role in working towards achieving much of this agenda and in measuring and tracking progress, particularly in connection with the 33 indicators directly connected to child and youth development. Of these, UNICEF is the proposed sole custodian for 10, and joint custodian for a further six ("Revised list," 2017). All 33 of these indicators were considered to answer the above mentioned research questions. (A detailed list of these 33 indicators can be found in **Appendix 1**).

Some of these indicators have been designated by the IAEG as Tier I, meaning that measurement is conceptually clear, there is an established methodology available and data is regularly produced by countries. For a number of these indicators, however, existing measurement methodologies are not regularly undertaken or no clear methodology or standards exist. Indicator 4.2.1, for example, which aims to track the proportion of children under five years of age who are developmentally on track in health, learning, and psychosocial well-being, is Tier III; there is no established methodology for this indicator. This paper is based on the premise that perceptions measurement is a useful addition to this not yet fully developed landscape of SDG indicators.

I.b. Key concepts and assumptions

The definition of 'Perceptions Data' as the "expressed views (opinions or perceptions) of people on a particular topic" has been adopted from Rodriguez Takeuchi, Hine and

Chavez (2015). This excludes objective data gathering (e.g. questions asked around certain life events, material possessions or other reported facts like income levels, voting behaviour, and frequency of usage of public transport). 'Perceptions Data' is not simply subjective data renamed. Rodriguez Takeuchi et al. (2015) note that 'expressed views' here primarily constitute perception and opinion gathering questions as choice can have different methodological problems. Additionally the authors are very clear in stating that 'people' are the general population. Perceptions Data thus excludes expert opinion, even though it may be about their opinions and perceptions (Rodriguez Takeuchi et al., 2015).

Given this definition, it is unsurprising that Perceptions Data has largely been measured by surveys to date although, barring the My World Survey, no measurement tool has expressed measuring perceptions of progress towards the SDGs as an explicit goal. The Oxford Poverty and Human Development Initiative (OPHI), for instance, have added on a 'subjective' element to their questionnaires (Rodriguez Takeuchi et al., 2015). Other examples of global surveys which have some perception-measuring questions include the 'Gallup World Poll' and the 'World Values Survey', both of which have been designed to cover a large number of countries with their standardized questionnaire on religion, government, ethics and well-being. These global surveys attempt to be nationally representative but have struggled to reach conflict areas, vulnerable populations and the homeless. There also exist regional surveys like the Latinobarometer, where the countries targeted are more limited but the questions could be more specific. Finally, several countries have successfully set up national data gathering instruments that

incidentally also measure people's views, such as the Social Weather Station in the Philippines (Alkire & Samman, 2014).

In order to develop key conceptual lessons in designing tools to measure perceptions, only those that incidentally measure perceptions across multiple countries (despite the significant potential adaptability of these tools to specific local contexts) were considered for analysis. Additionally, to avoid scope creep and ensure that the findings of this paper are as useful as possible, the focus was on ensuring that these tools have sufficient general characteristics. Country-specific practical challenges, like the lack of existing methodologies and coverage to collect information on the 33 SDG indicators in say a sparsely populated wealthy arctic country compared to developing dense tropical one are therefore not considered in this paper's findings; the focus being on understanding how these tools could be best used in theory. Finally, only perceptions questions asked since 2014 (a year before the commencement of the Agenda for Sustainable Development) were analysed, or in the case where a post-2015 round is not publicly available, then the latest available was considered to keep the findings relevant and timely.

II. Literature Review

Building on the context outlined above, a review of the literature was undertaken to determine why measuring and understanding people's perceptions is important, how it can be applied to the 2030 Agenda for Sustainable Development, what lessons could be learnt from current tools incidentally measuring these perceptions already and what potential issues may be faced in this process.

II.a. Why is 'Perceptions Data' important for SDG measurement?

It has been argued that Perceptions Data forms a bridge between knowledge generated by Professional Social Enquiry (PSI) and ordinary knowledge. According to Lindblom and Cohen (1979), PSI consists of activities that focus on establishing a causal explanation of social problems and actions whereas ordinary knowledge is that which is generated by common sense, speculation, thought and casual empiricism. There remains an important need to bridge these two forms of knowledge, since the commonly used method of ignoring ordinary knowledge and preferring PSI for social problem solving often leads to inefficient decision making and prioritization (Lindblom & Cohen, 1979).

More importantly, effective Perceptions Data measurement and utilization in policy making would bring the field of Public Policy closer to Lasswell's original image - a multidisciplinary, contextual, normative and problem-oriented focus to decision-making

(DeLeon, 2006). Asking citizens affected by a particular policy to share their viewpoint ensures that rationalistic policy making is not simply used to mask a particular ideology. Doing so in a manner that emphasizes political choice and citizen participation would further lead to more insightful policy definition, making, analysis, and monitoring.

Examples of this need to reform the field Public Policy abound. The OECD Better Life Initiative has argued that despite the money spent on monitoring areas, traditional data collection have failed to indicate key events such as the 2011 Arab Spring (OECD, 2013). Indeed, traditional indicators such as GDP per capita were growing in key countries in the region. Gallup concurs, claiming that not enough focus was put on understanding how people in that region were thinking, and this lack of understanding contributed to the crisis (Gallup, 2012, cited in Rodriguez Takeuchi et al., 2015). This conclusion is supported by Stiglitz, Sen and Fitoussi (2009) finding that only one third of people of in the United Kingdom and France trusted official statistics (Stiglitz, Sen & Fitoussi, 2009). In addition, Leo (2013) found that only 16% of US development assistance in Africa went to any of the top three priorities of local citizens (Leo, 2013).

Rodriguez Takeuchi et al. (2015) also argue that Perceptions Data could potentially fill a gap in the timeliness and availability of data. Indeed, Groves et al. (2009) highlight the Consumer Sentiment Survey's ability to gauge American unemployment rates more quickly than other objective surveys using questions about perceptions of the American labour market. In contrast, data included in the MDG indicators report were as much as four years old, making it difficult to assess progress (Sharma, 2014).

Barrett (2010) points out that the UN's Food and Agriculture Organisation food insecurity scale has proved to be a fairly reliable indicator of potential malnutrition crises and is being used as somewhat of an early warning system to target vulnerable populations (Barrett, 2010).

Additionally, one could argue that for certain SDGs, such as goal number five ("Gender Equality"), a key element of the goal is for the evolution of social norms, which can be challenging to measure using only objective data. Perceptions Data can be of particular use in understanding the beliefs that underlie certain taboos. The OECD's Social Institutions and Gender index is an effort to blend measurement of the objective legal framework, with Perceptions Data on the underlying social norms. Measuring these social norms and perceptions not only allows for a more detailed understanding of the de facto situation, but enables policy makers to better plan interventions (Rodriguez Takeuchi et al., 2015).

Price (2007) noted that while traditional engagement of the public has taken place at the evaluation phase of policy making and execution, a complementary shift to public involvement at the problem definition phase, where they can share their thoughts on what policy areas need the most attention and what they would like the output of policy decisions to be, would better play to their capabilities (Price, 2007). This directly links to the current motivation behind using Perceptions Data to not only create but also measure progress towards the SDGs. An example of using Perceptions Data for problem definition was the global 'My World Survey' which asked citizens what was the

most pressing social concern for them and their families in order to inform the crafting of the SDGs (UN News Centre, 2014).

Finally, this proposed use of Perceptions Data bears some similarities to the emerging literature around policy crowdsourcing, which is also being enabled by new communications technologies. As alluded to above, this has been proposed mainly at early stages of the policy cycle, however Prpić et al. (2015) argue that utilisation of online platforms for virtual labour markets, crowdsourcing tournaments, and open collaboration can be utilised at every stage of the policy process. They discuss primarily how existing platforms can be utilised to solve problems, generate ideas, and leverage the dispersed knowledge of groups. While the purpose, and goals of these crowdsourcing approaches are different to that covered in this paper, the review of relevant literature found many of the same advantages and drawbacks as outlined in the following section (Prpić et al., 2015).

II.b. How have current tools been measuring perceptions and what lessons have been learnt?

Rodriguez Takeuchi et al. (2015) highlight that while representation is worth monitoring, the strength of Perceptions Data is its frequency and timeliness in gauging when policy intervention is required. This strength relies on the fact that what opt-in modes of questionnaire delivery like SMS- and web-based surveys may forego in representative sampling, they make up for this in speed of response (Rodriguez Takeuchi et al., 2015). In short, while there is some division in the literature on the ideal

method of measuring Perceptions Data in relation to the SDGs, the technology used in this measurement is a central factor. Given the rapid pace of change in technology in the last three decades (i.e. the spread of mobile and smartphone technology and the advent of the internet), it is not surprising that there has been no systematic review of which technologies have been, or are currently being, used for gathering Perceptions Data related to SDGs (at the time of research, the UN Statistics Commission is examining the use of Perceptions Data in goal measurement).[1] It is possible then, that the usage of Perceptions Data to help measure SDGs will favour certain data collection methods over others that are usually recommended for subjective data (e.g. opinion polls, subjective well-being surveys etc).

On the other hand, Rodriguez Takeuchi et al. (2015) point out that main method of Perceptions Data gathering has been surveys with the mode depending on coverage. The most direct use of Perceptions Data in relation to SDGs to date is the *MY World* Survey as discussed earlier. Howard and Wheeler (2015) have argued that the survey's technology limited its usefulness as it relied mostly on paper ballots distributed by partner NGOs which may not have been representative (it did also use web and SMS ballots, though so far these two modes combined have accounted for a little over one million of the 9 million responses) (Howard & Wheeler, 2015). A few other indirect tools that frequently measure perceptions in their questionnaires are the Gallup World Poll and the World Values Survey. While the Gallup World Poll splits modes relying on telephone interviews where coverage is sufficient (mostly developed countries) and

[1] The 2017 UN World Data Forum (held in January), for example, hosted a plenary session by Afrobarometer on the use of Perceptions Data in relation to the SDGs.

face-to-face (FTF) interviews, the World Values Survey has used telephone interviews in special circumstances where representative coverage would otherwise be difficult to achieve (World Values Survey, 2016). In short, so far major global tools have mostly relied on more traditional survey technology.

Nonetheless there is a growing body of literature on using information communication technology (ICT) to conduct opinion surveys. Lüge (2015) reviews eight SMS platforms all of which are capable of two-way communication and points out that currently, SMS-based data collection will reach a wider audience than instant messaging (which relies on smartphones or notebook access) (Lüge, 2015). Interestingly UNICEF Myanmar has had some success at building a group of U-Reporters utilising only Facebook, however this has only occurred in the last 12 months which explains why no systematic reviews related to this approach have been found. Leo and Morello (2015) argue that in practice subjective poll data is better gathered via a call center rather than by Interactive Voice Recognition (automated software) (Leo & Morello, 2015). Both authors note that text-based questionnaires rely on high levels of literacy to be representative.

One significant advantage that new technology has the potential to bring is that of reduced cost. While Bamberger et al. (2016) acknowledge that the initial costs of establishing IT systems in terms of both infrastructure and software development can be high, they enable significant economies of scale that can lead to cheaper measurement. Additionally, Morrow et al. (2016) outline how the World Food Programme's mobile Vulnerability Analysis and Mapping (mVAM) was able to more

effectively understand the needs of particularly vulnerable groups using technology.[2] Furthermore, mVAM was also able to deploy this technology quickly during the 2014 Ebola epidemic in West Africa where it may have been unsafe to respond using traditional methods. By utilising new technologies these groups were reached while reducing costs from $20-$40 for FTF surveys to $5-$9 ($5-$6 for SMS alone) using a mix of phone calls, Interactive Voice Response (IVR) and SMS (Morrow et al., 2016, World Food Programme, n.d).

Additionally, when first rolling out mVAM a number of different approaches were tried to build a better understanding of the biases outlined above, including distributing mobile phones to vulnerable communities in refugee camps, and utilising mixed methods approaches combining traditional FTF surveying with new technologies. They found that attrition of the sampling frame was a significant challenge, and that food insecurity was higher despite a probable bias to wealthier and literate households. A key lesson acknowledged by Morrow et al. (2016) is that survey design needs to be iterative, and questionnaires must be piloted, indeed without the enumerator clarifying or providing context clues, ensuring the questions are clear to respondents is of even greater importance (Morrow et al., 2016).

II.c. What issues have been identified in utilising these tools?

The main areas of concern when gathering Perceptions Data is balancing between the classical 'objective-subjective' divide when it comes to social science research. As

[2] These groups often live in areas isolated by conflict or poor infrastructure, making data collection costly and potentially dangerous

highlighted above and reiterated by Veenhoven (2004), there is an urgent need to bridge this divide and collectively use both sources of data and insights to inform policy decisions (Veenhoven, 2004). Some of the key criticisms of subjective/Perceptions Data that has prevented this collective usage from taking place is the concern that the data collected is not reliable, representative of all citizens nor does it include populations living in inaccessible regions. (Herbert, 2013).

Additionally, there is the issue of existing subjective data being ignored by policy makers. Seaford (2013) attributes this to the fact that subjective indicators such as perceptions indicators do not normally take centre stage in public debates and thus politicians are not held accountable to them by the public. It is further suggested that the reason for this not happening is that subjective indicators are often not as simple and easily understandable as macro-level objective indicators like GDP. Their analysis is also not focussed on identifying key drivers and causal pathways, making it difficult to tangibly interpret the data in ways that could prove useful for policy. This means that people are unable to make a decision on whether the subjective indicator matters to them or not, leading them to not hold politicians accountable to them, although this could be changing (Seaford, 2013).

Bamberger et al. (2016) argue that the focus on the ability of these new technologies to generate and analyse data far quicker, and at significantly lower cost than traditional methods has led to less focus on issues such as data quality, relevance and selection bias than should be the case. With average adult mobile penetration rates in Africa of 73% and growing, there is certainly the potential to reach large numbers of people

relatively cheaply, even in developing contexts. However, the ability to achieve a degree of representativeness is dependent on high penetration rates, with the internal distribution of who has mobile phones also important. For example, the Centre for Global Development and the World Bank attempted to build nationally representative samples using mobile network operator data. While this was relatively successful in Zimbabwe, in Afghanistan mobile penetration is much lower, and a number of factors have restricted the adoption for rural women, so representativeness has been much harder to achieve (Bamberger et al., 2016).

Bamberger et al. (2016) also point out that users of these new technologies need to confront is the fallacy of large numbers. While crowdsourcing and other techniques enable large sample size, the skew in access towards wealthier and more literate individual means that the results may not reflect the broader priorities and thoughts of the community. One key way to combat this is through defining the sampling frame first, and directly providing mobile phones when necessary, to the population of interest (Bamberger et al., 2016).

Comparing costs is not as simple an issue as it might seem. The World Bank Guide for Mobile Phone Panel Surveys (MPPS), for example, outlines three major contributors to the cost of a panel survey: conducting a baseline survey; the costs of hardware including survey infrastructure and the mobile phones; and a call centre. Yet for an SMS or app based approach, many of these costs can be reduced, particularly per data point as they reach a scale of respondents, although when the costs of SMS are included it can be difficult to be cost competitive compared to the very comprehensive surveys

that can be run. They caution that comparing costs with traditional FTF surveys is likely to be misleading, and lead to different conclusions in different circumstances, with mobile surveys able to respond rapidly to changes, and collate and disseminate data much quicker (Dabalen et al., 2016). That being said, mobile phone operators sometimes bear or donate the costs of SMS, and in these instances, there can be a significant cost advantage.

Finally, there is the problem of conflicting and overlapping worlds of public, private and civil society data, in a time where both citizens and policymakers in parts of the world are facing the problem of too much information rather than too little. According to the Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAGDRSG) (2014), here are rapid technological advancements that lend themselves to innovations in this field, but there is also a lack of higher level coordination and mobilization of the collected data, leading to multiplicity, inefficient use of resources and issues in securing sustainable funding (IEAGDRSG, 2014).

III. Analytical Framework

Given the lack of literature available, this paper has designed an Analytical Framework to establish what would make a tool effective in measuring perceptions, drawing from established survey methodology guidelines, preliminary research and expert consultations. Groves et al. (2011) have mentioned three criteria that all data measurement tools should meet: is the right thing being asked, do the respondents understand the questions being asked and can they (as well as the interviewers) complete the questionnaire easily. Building on these standards, they specify certain guidelines to measure people's attitudes that can be applied in the context of Perceptions Data Measurement. These include:

- Specifying the perception object being measured clearly in the question;
- Measuring the strength of this perception by providing multiple options if needed;
- Using closed questions to ensure comparison of data, preferably with five to seven point scales, with each point being clearly labelled;
- Avoiding double-barreled questions, leading questions and alternatives that would influence the answers of the respondents; and
- Collecting data on the same questions over different periods of time in order to measure change in perceptions (Groves et al., 2011).

These guidelines, however, were created keeping in mind traditional data collection methods like long-form surveys as opposed to U-Report's SMS and instant messaging

platform. To further provide for technological advancements like Real Time Monitoring (RTM), this paper uses additional preliminary research. Chai and Cummins (2014) highlight the role RTM can play in linking data collection to a relevant and timely response mechanism as well as in increasing the frequency of measurement (Chai & Cummins, 2014). Building on this theme of timeliness, Rodriguez Takeuchi et al. (2015) emphasize the importance of maintaining the reliability of the perceptions being gathered so as to draw comparisons across time, countries and different groups of people within those countries (Rodriguez Takeuchi et al., 2015).

This knowledge base helped formulate a first iteration of the Analytical Framework which was subsequently revised based on consultations with academics and development practitioners. The final version then defines four key standards to determine what an effective measurement of perceptions would be, as mentioned in Figure 1. A brief outline of the rationale for the inclusion of each element is mentioned below (see **Appendix 2** for more details).

While the cost of utilising these tools to measure perceptions is clearly an important consideration, practitioners interviewed by the research team were unwilling to publicly reveal the costs of data collection. They have, however, given general guidance in confidence. It is understood that the scale of costs savings is often similar to that of WFP's mVAM. It should be noted that mVAM does not rely on a single platform, and as such the economies of scale may be smaller than could be achieved for a tool such as U-Report, particularly as it is now an established platform.

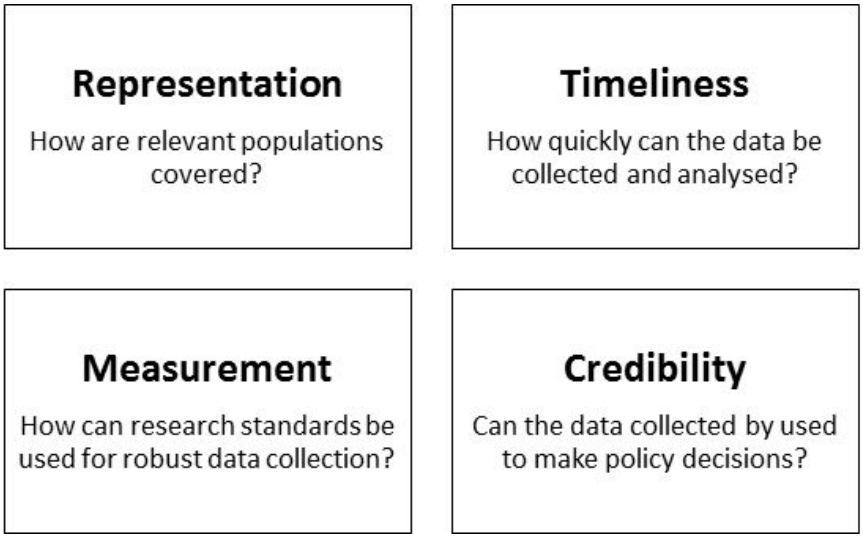


Figure 1: Analytical Framework to evaluate the measurement of perceptions

III.a. Representation

Most survey literature identifies representation as a central factor for judging the effectiveness of a tool (ESOMAR/WAPOR 2014, Link et al., 2014). In this Analytical Framework, the Representation standard captures whether a tool measuring perceptions covers all the populations relevant to its focus area of data collection, has a robust sampling frame and design and provides the metadata needed to evaluate its coverage (Groves et al., 2011). As discussed by Rodriguez Takeuchi et al. (2015), capturing, measuring perceptions for the SDGs is helped by identifying the views of specific groups in a timely fashion (Rodriguez Takeuchi et al., 2015).

III.b. Timeliness

As outlined in the Literature Review, the lack of timeliness of the data available to measure SDGs in a number of countries is a key challenge, with data in the monitoring

report up to four years old (Rodriguez Takeuchi et al., 2015). New technologies for measuring Perceptions Data can potentially address part of this gap in understanding and provide initial indications of the success of government, donor and lender interventions. Timeliness here is therefore the ability of a tool to quickly collect and analyse data to generate usable results. Firstly, this includes the frequency of data collection and the time spent on processing the results. Chai and Cummins (2014) highlight how RTMs aid in better targeting policy interventions, speeding up local service delivery, strengthening planning for crises due to a tighter cycle between data collection and the reporting of results as well as building response capability (Chai & Cummins, 2014). Secondly, there is a focus on the continuity afforded by tools with Groves et al. (2009) pointing to the importance of asking the same questions to particular population groups so that more robust comparisons can be made across time (Groves et al., 2009).

III.c. Measurement

The third key standard in the Analytical Framework is Measurement, which evaluates for potential recall biases (the ability of people to effectively remember past perceptions or feelings), social desirability biases (whether the question invites an individual to admit to illegal behaviour or criticise the government), framing effects (does the question lead the respondent to a particular answer?), and whether the respondent is a proxy (i.e. are they asked to answer questions about other people's behaviour?). If any of these measurement issues are found then they need to be kept in mind by any policy maker

seeking to use the collected Perceptions Data (Rodriguez Takeuchi et al., 2015). Additionally, drawing from the guidelines laid out by Groves et al. (2009), the framework evaluates the wording of the questions (whether a single object is being measured and is state clearly) and whether the strength of the perception is being measured (using closed questions and labelled scales) to ensure the data collected is robust and comparable.

III.d. Credibility

Credibility becomes a key evaluation criteria given the focus on using Perceptions Data on maximising country progress towards the SDGs. Peixoto and Fox (2016) highlight the importance of the institutional and technical design of the tools so that citizen concerns and the relevant government service providers can be identified precisely. Therefore, a tool that effectively routes citizen perceptions to the relevant providers within governments will ensure a stronger uptake of feedback. Also, publically disclosing the given feedback will demand a higher accountability from countries when it comes to implementation (Peixoto & Fox, 2016). Based on this, the Analytical Framework gauges the country context provided by the tools that measure perceptions and the easy availability of the results for decision-makers.

To summarize, this Analytical Framework has been designed to cover how effectively are perceptions being measured by current tools, not only from the theoretical and research-based aspect of robust Representation and Measurement but also to include the practical applicability of the results and its demonstration of Timeliness and

Credibility. The next section defines our methodology of applying this framework to key existing tools so that interesting practices and recommendations can be drawn out to improve further usage of Perceptions Data in meeting the SDGs.

IV. Research Methodology

While the Analytical Framework lays out a theoretical understanding of what the effective measurement of perceptions would look like, to analyze current tools for their effectiveness and ability to contribute to the monitoring of SDG progress, a modified Process Benchmarking has been adopted. The RAND Corporation defines Benchmarking as a method to compare performance or processes across organisations, policies or systems with the aim of generating ideas of what improvements can be made, which complements the aims of this paper (Ling & Villalba van Dijk, 2009). Process Benchmarking is a multi-stage process involving defining what is to be benchmarked, identifying the participants to be benchmarked, gathering data, analysing it to generate recommendations, executing these plans and finally monitoring their impact on outcomes. For this paper the focus will be on conducting the Benchmarking study up until the laying out of recommendations and an implementation plan. The actual execution of the changes suggested and its subsequent monitoring is not possible given limited resources and time. However, the hope is that the findings from this paper will be used to inform key changes in the measurement of perceptions to increase the efficiency of monitoring progress towards the SDGs. The method of conducting the Process Benchmarking have been adapted to create a three stage methodology for this paper discussed below:

IV. a. Stage I- Identifying tool that measure perceptions and collecting data

For tools to be considered in this paper, they would need to meet certain inclusion criteria that would ensure the generation of relevant data to evaluate using the Analytical Framework. These criteria are that the tool:

- Has some questions that currently measure perceptions;
- Measures people's perceptions on the target Indicators for this paper;
- Has results available, disaggregated by key demographic factors; and
- Has publically accessible research methods, questionnaires and results databases.

Given the emerging nature of the term 'Perceptions Data' and the lack of literature around its effective measurement, the list of tools to consider was built by initially conducting a keyword search on a variety of search engines. This was narrowed down based on the inclusion criteria established above, expert consultations and literature reviews. Finally, tools only covering single countries were excluded to ensure the applicability of any findings across contexts. An exception was made for U-Report, where three separate country tools were considered to make-up for the national identity of this tool. Specifically, Nigeria and Uganda were considered because they have been in existence for the longest and have the largest user base. Myanmar was considered to represent U-Report's presence in a different cultural context in Asia and since it is the most recently launched tool with usable data available. It has also relied on Facebook instead of SMS, and therefore provides an interesting alternate use case. This led to the

identification of the following tools that measure Perceptions (further details of these tools can be found in **Appendix 3**).

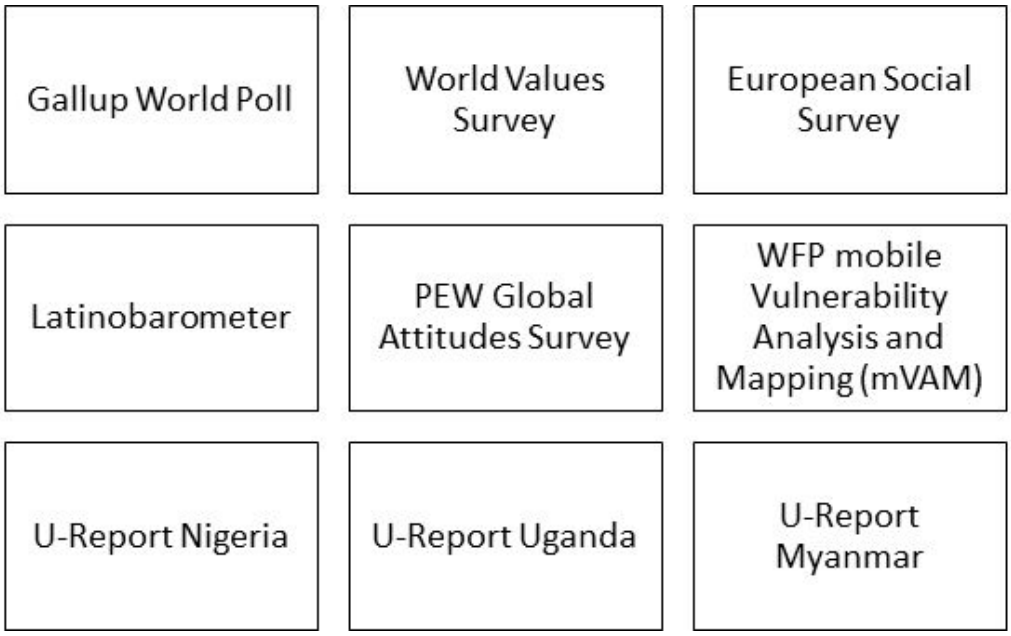


Figure 2**:** List of Tools evaluated using the Analytical Framework

Once the tools were identified, a majority of the relevant data for evaluation was collected from the webpages of these tools, with individual interviews with the relevant managers added if necessary, to fill the knowledge gaps. Due to the technical criteria included in the Measurement Standard, this process was further built upon with the following steps:

- As a research team, each criteria under the Measurement Standard was clearly defined based on standard survey methodology guidelines with example questions generated that met and didn't meet the criteria;
- Each researcher owned the data collection for three tools and double-checked the data for three others;

- Researcher I identified questions in the tool that measured people's perceptions based on the definition outlined above and linked them to the relevant SDG goal, target or indicator;
- Researcher I then identified the measurement characteristics of each question based on evaluation criteria laid down in the Analytical Framework. Subsequently the results were averaged out to generate one data point for each criteria under the Measurement Standard; and
- Finally Researcher II double-checked the interpretation of the question's characteristics to identify any possible deviations from the criteria definitions agreed upon. Any difference in opinion were addressed by Researcher III.

IV. b. Stage II: Evaluating the data and applying the findings

Once the relevant data was collected, key differences and similarities between the various tools were identified. This was done by evaluating each tool against the criteria of the Framework to identify strengths and deficiencies that exist in their current way of operating, that may aid or hamper the collection of Perceptions Data for the SDGs (further details of the method of evaluation and actual results are found in **Appendices 4-7**). These findings are elaborated on in Section V of this report.

IV.c. Stage III: Generating recommendations

Finally, the findings were used to suggest changes that could be made to existing tools measuring perceptions (especially the U-Reports), to make them more effective in

monitoring SDG progress. Three key recommendations were generated which include implementation plans that not only overcome current gaps in the monitoring process but also make the field of perceptions measurement more effective. These recommendations are elaborated on in Section VI of this report.

V. Findings and Analysis

The data collected on the four standards of the Analytical Framework was firstly used to identify similarities and differences between the tools measuring perceptions and factors contributing to these differences. This analysis interestingly identified four key findings: firstly, two typologies of tools can be identified showcasing the trade-off between the Representation and Timeliness Standards; secondly, while tools designed for representation were more direct and controlled in their measurement of perceptions, biases still remained across both typologies; thirdly, information on the questions and tool methodologies was generally open though both the U-Reports and Gallup were more complex to access; this is slightly unfortunate as our fourth finding is that these four tools have covered the most applicable SDG content in their perceptions measurements.

V.a. Finding I: The two typologies of tools measuring perceptions

Analysis of the nine tools leads to the identification of two types; the first, (the World Values Survey (WVS), European Social Survey (ESS), Gallup World Poll (GWP), Global Attitudes Survey (GAS) and Latinobarometer) focusses more on representation whereas the second type (the three U-Reports and mVAM), has timeliness as its strength. As outlined in the Analytical Framework both of these are key for gathering perceptions but amongst the tools surveyed, achieving both has been a challenge for the following

reasons (see **Appendix 5** for detailed breakdown of the evaluation on the Representation and Timeliness Standards).

Considering the tools with timeliness as their strength, they are able to issue multiple polls with results shown within 24 hours, and often in real time. Unlike the representation heavy surveys, the same populations can be surveyed (although the same respondents are not guaranteed). Perceptions in the case of U-Report are ongoing and easily updated. Opinions of U-Reporters can be gathered when they are needed almost without delay. This comes with a cost in representation, both nationally and to a certain extent, in terms of populations targeted by SDGs. While additional information does allow for some representation (location for example), U-Report currently relies on respondents to self-select, be literate and have access to a phone to provide such fast responses. Even in Nigeria where only 11% of the population do not own a phone, this could leave out groups most in need of the target SDGs (e.g. the 40% of the population that cannot read) (Pew Global, 2015; CIA World Factbook, 2016). Aggregation, according to interviews with U-Report staff is possible, but based on voluntary information that is sometimes left blank. U-Report Myanmar relies on Messenger but in our tests did not seem to verify the respondent's answers concerning their identity, with the Facebook profile. Therefore the U-Report model works for its current purpose of information dissemination and creating youth influencers but may lead to roadblocks when it comes to monitoring progress towards the SDGs.

For the tools with representation as their strength, achieving this hinges on responsiveness and aggregation. This currently comes at the cost of timeliness and

may still fail to cover vulnerable groups. Tools that scored high on representation did so partially through large sample size requirements. WVS for example requires at least 1,000 respondents in each country surveyed while the ESS sets this at 1,500 for all countries with populations above 2 million. Gallup, WVS, and the ESS all rely on randomized selections of longstanding widespread registries (household registries). To date, guaranteeing such large quotas of respondents has meant that these representational tools have preferred face-to-face interviews, using telephones only where there are nationally representative lists of phone numbers. Gallup, for example, requires 80% coverage by such a list. While conducting telephone and computer assisted face-to-face interviews (CAPI) gives control and access to those who cannot read or do not have access to cell phones, they require the ownership of a phone and a stationary address. In all of these, the costs of not just time, but also money were much larger than those of U-Report.

Timeliness of the more representative tools was also much more limited with the most rapid turnaround between data intake and published results being Latinobarometer which (unofficially) takes seven months. GWP is conducted the most often out of these tools, but this is at most semi-annually and varies from country to country. Question formation in most cases takes time as well. ESS for example takes 20 months to form and test questions to ensure cross-country compatability.

In short, none of the tools analysed are ideal both in representation and timeliness. Quickly generating questions about perceptions and gathering responses is low cost but creates a selection bias and limits whose perceptions are being gather. Ensuring

representational responsiveness that bypasses problems of literacy with standardized tested questions takes time.

V.b. Finding II: Open anonymity and controlled precision in measurement

Our findings show that there are also no existing tools that exclusively use ideal perceptions questions though weaknesses and strengths vary across the tools (see **Appendix 6** for detailed breakdown of the evaluation on the Measurement Standard). The quality of these questions in terms of their ability to limit biases varied. None of the tools measured suffered heavily from recall biases, and the variance did not correlate with the divide between timeliness and representation. The ESS and Latinobarometer had the largest proportion of questions potentially biased by recall at around 18% of questions asking about past feelings (ESS) and 11% (Latinobarometer) asking about past perceptions. U-Report Nigeria stands out for having none despite asking the largest number of perception questions.

Perceptions questions that evoked a social desirability bias were more frequent especially in discussing the workings of the government, with 59% of Pew's questions utilizing this content. While the literature suggests that such questions create a bias, this needs to be considered in context with the anonymity provided by the tool. Pew uses face-to-face interviews to collect surveys which forces people to admit to another person negative perceptions of their government. Actual measurement of the bias this creates is beyond the scope of this evaluation, nonetheless, a high proportion of perceptions questions about government activity may be less desirable without

anonymity. Discussing illegal activity is similarly dependent on this context. With the exception of Pew which stands out for having no such questions, between a third (WVS and U-Report Uganda) and a fifth of questions (ESS and U-Report Nigeria) in the tools were about activities which may be illegal in the countries surveyed. While the WVS uses terms like 'justified' to discuss illegal activities (for example, parents beating children) in person, none of U-report Uganda's questions implied the respondent was undertaking illegal activity. The fact that the latter is anonymous also means that the social desirability bias here is much less likely to be a problem (Groves et al., 2009).

Leading questions, on the other hand, were more common in the timely tools with Gallup being the exception (at 22% of questions being leading). U-Report Uganda, for instance, had more leading perceptions questions than not (66%), in large part due to the tool's dual role as both an information gathering and delivery tool. A typical U-Report Uganda leading question delivers news before asking a related question (e.g. "It's important dat a pregnant women go 2 da 1st antenatal care (ANC) visit 2 learn abt danger signs but only 32% go. Why don't more women go 2 da 1st ANC visit?"). While this may increase responsiveness, a cursory analysis found no significant correlation in this tool between the two factors. This calls into question whether the potential framing bias should be reduced when measuring perceptions to monitor SDG progress.

Perceptions questions about other people's perceptions (for example, the Gallup question, "*do all members of your household get enough food everyday?*[3]) were between

[3] Others in that household might not agree that they get enough food everyday. See the *Analytical Framework* for more details.

a tenth and a quarter of questions for most tools, with the exception of ESS and Pew which had none. The U-Reports tended to have the highest proportion (Uganda had a ratio of 26% while Myanmar and Nigeria both had 16%). The more representational tools had slightly lower proportions (only 7% of Gallup's questions had this type of measurement issue) suggesting that this may be according to typology.

Clearly stating the perceptions object in the question, on the other hand, is necessary in perceptions. In all tools, the majority of questions did this, though U-Report Uganda and Latinobarometer having the lowest ratio with 28% and 25% of their questions featuring questions in which the object could be confused. The question "*U-report: How can businesses invest, to enable U to become the best employee/most productive u can ever get? We will share ur replies with the business sector*," for instance, does not require the respondent to say whether his answer is reflecting either investments towards being the best employee ("best" is also undefined) or most productive. Here there was no clear divide between the timely and representational tools with no tool completely avoiding this type of question (though U-Report Myanmar did have the lowest proportion at 5%).

In terms of response collection, almost all questions in the representational tools were closed (WVS and Gallup, only used closed questions while 90% or more were closed in ESS, Pew and Latinobarometer). With the exception of Gallup, the strength of the perceptions were also usually measured in the more representational tools (in Gallup only 35% did so). For timely tool questions there were more open questions (20% in U-Report Nigeria and 60% in Uganda and Myanmar). Unsurprisingly, direct measurement of perceptions strength (for example in the Gallup question "*In your opinion, how serious*

is the problem of air pollution where you live - very serious, somewhat serious, not very serious, not serious or not at all serious?") was also less frequent. Timely tools tended to favor either keyword measurement (U-Report Uganda and Myanmar) or labelled, but non-scaled options (e.g. "why do you think people practice open defecation?"; "people do not care, people have to wait too long on a qu, the latrines or toilets are dirty and, people have no latrines or toilets") in the case of U-Report Nigeria. This divide in the measurement of the strength of perceptions ties into the fact that more representational tools labelled their scaled strength questions (95% of Pew's for example) in comparison to U-Reports who labelled in around one third of cases (U-Report Nigeria), or less than one in 20 (U-Report Myanmar and Uganda). This is potentially due to the constraints of SMS technology, where listing a detailed scale can add to the length of the message significantly.

In sum most tools had some issues with recall bias with Social desirability bias also prevalent, though theoretically less of a problem in the U-Reports due to their reliance on anonymous SMS communication. On the other hand, leading questions and unclear objects were more frequent for these more timely tools, calling for certain adjustments to be made before perceptions towards monitoring SDG progress can be measured effectively.

V.c. Finding III: Accessibility of Data

Although our criteria for inclusion of a tool was that it had accessible data, levels of accessibility was not uniform with none of the tools explicitly dividing subjective from

objective questions. In the case of U-Report, questions are not searchable by time and are sometimes difficult to find within a flow, or set of questions. U-Report Nigeria does have a filter for SDGs, however, it currently does not return any previous polls. In the case of Gallup, question lists are released but the results are only accessible after payment. In contrast the ESS releases all results and questionnaires (with answers) in multiple formats for public download, though as noted before, this is done every two years.

What stands out then is that the tools were all relatively accessible in how they gathered perceptions (see “Finding I” and **Appendix 7** for a detailed analysis on the Credibility Standard). Questionnaires were adapted linguistically and culturally with issues arising instead in terms of accessibility to results and no ideal tool practice emerging (though ESS’s multiple data formats does provide a standard in terms of downloadable data). A key strength that stands out here for the U-Reports is their ability to make results actively accessible to key decisions makers. This means that not only are the results of their polls published online (within a day of the answers being collected) but are also actively fed back to key CSOs and Government authorities via the advocacy efforts of UNICEF. This results in timely action being taken on issues of importance as is highlighted by a key success story from U-Report Nigeria as gathered from an interview with the country representative. A poll on the importance of providing cheap Ready-to-use Therapeutic Food (RUTF) to solve child malnutrition was run in June 2016 with the results being used by UNICEF to appeal to the wife of the President. The

campaign successfully led to the Government allocating spending for RUTF within the 2017 Nigerian Budget for the first time.

V.d. Finding IV: Four standouts for SDG coverage

Finally, the most applicable measurement tools in terms of perceptions questions related to SDGs were the U-Reports and Gallup World Poll. U-Reports Uganda and Nigeria stand out in terms of having the widest coverage of SDGs (four were focussed on, see **Appendix 8** for a detailed mapping of tools to the SDGs covered - each number represents the number of questions considered). This is perhaps not surprising given their volume of perceptions questions, though it is worth mentioning that Gallup also focussed on four goals. U-Report Uganda tended to be less precise - more questions related to goals rather than to specific targets or indicators. U-Reports Nigeria and Myanmar and Gallup, by contrast, asked more questions related to specific indicators (roughly a fifth in the case of U-Report Nigeria and over a quarter in the case of U-Report Myanmar and Gallup). Note that this does not mean other tools could not adjust their questionnaire content more towards SDGs in the future. This finding is relevant in that it gives a picture of what perceptions tools are available right now in terms of content to cover the selected SDGs.

VI. Recommendations

Building on the Findings above, the recommendations answer the second part of the research question: How can these tools for measuring perceptions - especially U-Report - be used to enhance the measuring and reporting of SDG progress?

VI.a. Confronting the trade-off between timeliness and representation

Currently none of the tools studied in the paper have been used for measurement of SDGs, with only WVS expressing a public desire to adapt to this purpose. The key question for all of these tools is essentially, can the underlying technology or approach be used to contribute to measurement in a manner that is useful? While there are a number of tools that have developed and continue to refine their approach in line with academic standards to ensure that they provide robust and valid information, some of them can be rolled out as irregularly as once every five years, and there can be years in between data collection and publication of results. The usefulness of the data so generated is clearly diminished through this lack of timeliness which is exacerbated by the potential for volatility in the data.

On the other hand, new ‘pulse’ style tools enable extremely rapid data collection and dissemination but are potentially subject to a greater number of biases as outlined above. These biases limit the usefulness of the data as they could lead to invalid yet influential generalizations forming in the long run, as referred to in Section II. However, new technologies allow more than just advantages in timely data, they could also

enable the greater measurement of vulnerable populations, who often live without infrastructure connectivity, or surrounded by conflict, and where data collection through traditional means is often time consuming, expensive, and potentially puts the enumerators at personal risk (Morrow et al., 2016).

How then to ensure that the pros and cons of each approach can be appropriately weighed? In particular, what role can these 'pulse' tools play in this measurement system, given the reliance to date on traditional measurement methods to gather Perceptions Data? One potential approach, which has been utilised by the World Food Programme with mVAM, is to examine the validity of new survey tools through adopting a mixed methods approach, combining SMS (or other technology based) short form surveys.

UNICEF should consider how the Multi Indicator Cluster Survey (MICS) could be integrated with U-Report (at least in the timing and alignment of some questions) in some countries in order to examine the validity of the sampling frame and understand the biases present. Given that MICS is currently under review to see how it can better report on some of the Tier III indicators, an opportunity exists to see how this could be effectively achieved. This could alleviate much of the requirements of undertaking a baseline survey outlined in Section II.b.

Furthermore, these emerging technologies enable significant advantages in cost and flexibility. Where partnerships exist, and sufficient privacy protections are in place, then UNICEF should also consider how a representative sample could be built from existing

U-Reporters, and supplemented where needed, building off the case studies outlined by Bamberger et al. (2016) in Section II.c. Where representativeness continues to be difficult to achieve, the distribution of mobile phones to target populations is something else that could be considered to improve responses, although this is likely to have an additional cost. UNICEF is already proceeding with a trial in Italy in providing mobile phones directly to new refugees in order to use the U-Report platform to best track the needs of this vulnerable group. Lessons from this trial and those of mVAM and the World Bank should be incorporated into doing this effectively, and it will likely vary substantially by the context of the target group and specific country. Additionally, UNICEF should consider how U-Report can integrate audio, voice, or IVR into the tool to reach those who may not have sufficient literacy to engage otherwise (Leo et al., 2015).

Finally U-Report has a large base of U-Reporters in some countries, with more than 2 million on the platform in Nigeria for example but response rates and engagement can vary widely. Feedback from interviews with U-Report managers in different countries suggested various reasons for non-response, including: interest in the topic (often influenced by what is topical in local news); understanding of what the information is being used for; and the relationship developed with the U-Reporter through previous information provision or hearing their views. These elements need to be kept in mind when attempting to measure SDGs as they could induce further bias in the sample. As U-Report continues to grow attrition only appears to be a problem in so far as it promotes bias. Continuing to ensure that U-Report is easy to use, and relevant to local populations will help reduce the psychological costs that lead to lower response rates,

and should keep individuals engaged with the platform. In more mature markets where U-Report has been operating for a number of years a more in depth analysis should be taken of its user base to see if there is any consistent profile of those who cease to use the platform, or only engage sparingly (Dabalen et al., 2016).

VI.b. Framework for effective question design

A key aspect of this paper is the application of the Analytical Framework developed to ensure that the questions utilised to gauge people's perceptions are well designed and that the gathered information is as useful as possible. Building on the findings in Section V, the key recommendations when seeking to design perceptions measurement questions are summarised below.

Perceptions object: What the question is asking needs to be clearly defined. If the question is not clear, or if it is double barrelled then it is likely to be difficult to effectively interpret the output to the question, reducing its utility.

Example question to consider: "In your opinion, how serious is the problem of air pollution where you live - very serious, somewhat serious, not very serious, not serious or not at all serious?"

Example question to avoid: "U-report: How can businesses invest, to enable U to become the best employee/most productive u can ever get? We will share ur replies with the business sector"

Social desirability bias: While it is sometimes impossible to avoid asking questions that relate directly to the working of the government, social taboos, or illegal behaviour, as that is the topic that is of interest, asking individuals to admit behaviour or perceptions that may have social repercussions should be avoided where possible. This is one potential advantage of computer or mobile device assisted methodologies as they don't require the admission of such behaviour to another human being. When this is the topic of interest, questions should be carefully crafted.

Example question to consider: "Thanks sharing, final question 2 u as a woman: What is the biggest issue u face during your periods? Ur answer is anonymous & free across all networks"

Example question to avoid: "Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card. For a man to beat his wife"

Leading questions: This was particularly an issue with the shorter form new technology surveys where quick background information was often provided prior to asking questions. With many of the example questions that were found as part of the research for this paper, this is likely to significantly bias the responses. It is possible that this information increases responses, but if response rates are a potential issue for these questions, then different types should be tested first with different sub-populations.

Example question to consider: "Hello U-Reporter, do you think Polio is still a health problem in Nigeria?"

Example question to avoid: "It's important dat a pregnant women go 2 da 1st antenatal care (ANC) visit 2 learn abt danger signs but only 32% go. Why don't more women go 2 da 1st ANC visit?"

Recall bias: This wasn't found to be a significant problem across the tools studied, however it is still important to note that generally people have a number of biases when trying to call past events, including over-weighting recent events, and priming.

Example question to consider: "Would you say you are very satisfied (1), fairly satisfied (2), not very satisfied (3), or not satisfied at all (4) with public hospitals or don't you have those services where you live (5)?" (Latinobarometer)

Example question to avoid: "Do you agree or disagree that this country has never enjoyed more social and political peace before?" (Gallup)

VI.c. Ensure tools contribute to enhancing monitoring of SDG progress

Despite the strengths of tools outlined earlier in this paper, they are not likely to replace traditional data gathering and measurement, and nor should they seek to do. These strengths ensure that they do have the potential to make a useful contribution to the broader measurement ecosystem. Firstly, many of the older tools examined in this paper have the data broadly available in a number of different formats for use of partners, CSOs, NGOs, public agencies and others. While U-Report should be

commended for the transparency of its data, with the results of all polls publicly available on the internet, it is not easy for researchers or others to aggregate those results, compare similar questions across different countries, or track the responses to similar questions over time. Therefore, some consideration should be given to setting it up in a way that makes it easy for the data to be interrogated, while maintaining privacy for individuals on the platform.

Furthermore, when seeking to utilise U-Report to measure SDGs some thought should be given to the creation of a portal for the range of data collected. New technologies, including sentiment analysis and big data, can be used in powerful ways to build a detailed understanding of progress at both a national and global level. The creation of an online open data portal that brings together all the data accessible to an agency such as UNICEF, ideally with some consistency applied to the formatting, would enable other organisations to use the data with significantly greater ease.

Finally, as UNICEF works with the IAEG to further develop the framework for global measurement of the SDGs it should consider how the input from Perceptions Data can be effectively paired with other forms of traditional data collected to form new indicators and indices. Bamberger et al. (2016) argue that greater use of ICT in evaluations of progress can lead to the oversimplification of complex constructs, and tools enable a greater focus on qualitative components to provide a richer story, and the temptation to reduce information to a single indicator should be resisted. An example of how Perceptions Data has been incorporated with traditional data gathering and other surveys is that of the CIVICUS Monitor. This tool integrates a range of different volatile

data sources and includes Perceptions Data, expert opinion, sentiment analysis, amongst other indicators to build a rating and a deeper understanding of the state of civic space in a particular country. Building something similar around key SDG indicators that UNICEF is custodian of, is beyond the scope of this paper, but is something that should be examined going forward. This would also enable this richer story to be more accessible to those seeking to understand progress on a target, goal, or indicator.

VII. Conclusion

This paper has attempted to understand how the lessons from the emerging field of Perceptions Data and measurement could be understood and applied towards effective measurement of the SDGs. A key conclusion is that this can be done, with the recommendations focussed on practical steps to ensure that the underlying principles and technology are applied in an effective manner. These recommendations can be summarised in the following three points:

- Confronting the trade-off between timeliness and representation;
- A framework for effective question design; and
- Ensuring that tools contribute to enhancing the monitoring of SDG progress.

While Perceptions Data was used in the design of the 2030 Agenda for Sustainable Development, there was no tool identified that is currently seeking to utilise Perceptions Data to measure these SDGs, although the World Value Survey are planning to incorporate these goals into their next Wave. This is despite the emergence of a range of different tools in recent years. In total nine tools were identified that have asked perceptions questions related to the SDG goals, indicators and targets that target youth. Based on the literature an Analytical Framework was developed, through which these questions and the tools themselves were evaluated. This detailed understanding of the status quo was then used to examine which of the underlying principles and approaches utilised by these tools could be effectively applied to SDG measurement

and reporting, so that these tools, with a particular focus on U-Report could be adapted to this purpose.

A key finding of the analysis is that there are two typologies of tools, with traditional long form survey instruments with representativeness as their strength being supplemented by the emergence of ‘pulse’ style tools that utilise new technology to reach large numbers of people, with the ability to scale and adapt rapidly. This leads to the first major recommendation of the report, confronting the trade-off between representation and timeliness. This includes greater use of mixed methodology approaches combining existing tools with new techniques to examine for biases, working with partners to build and test representative sample frames, and utilising clever design to address literacy, non-responses and attrition within the sampling frame.

In building an Analytical Framework capable of assessing what makes an effective tool, a significant number of questions were analysed, providing clear examples and a framework for how to gather perceptions on youth related targets and indicators. This framework is designed to help practitioners confront framing and recall effects, social desirability and recall biases, and other potential pitfalls of question development and design.

Finally, these tools are unlikely to replace existing measurement strategies and techniques, and nor should they. As outlined throughout the paper these tools have a number of strengths that can contribute to the existing measurement ecosystem. This involves bringing data collected into a single, easy to access open data portal, and

examining how different data sets can be effectively combined to build deeper understanding of the complex constructs that are the focus of the Agenda for Sustainable Development.

Abbreviations

ANC	- Ante Natal Care
CAPI	- Computer-Assisted Personal Interviewing
CSO	- Civil Service Organisation
ESS	- European Social Survey
FTF	- Face-to-Face
GAS	- Global Attitudes Survey
GDP	- Gross Domestic Product
GWP	- Gallup World Poll
IAEG	- Inter-agency and Expert Group
IAEGDRSG	- Independent Expert Advisory Group on a Data Revolution for Sustainable Development
ICT	- Information and Communications Technology
IVR	- Interactive Voice Response
MDG	- Millennium Development Goal
MICS	- Multi Indicator Cluster Survey
MPPS	- Mobile Phone Panel Surveys
MVAM	- Mobile Vulnerability Analysis and Monitoring
NGO	- Non-Government Organisation
OECD	- Organisation for Economic Cooperation and Development
OPHI	- Oxford Poverty and Human Development Institute
PSI	- Professional Social Inquiry
RTM	- Real Time Monitoring
RUTF	- Ready-to-use Therapeutic Food
SDG	- Sustainable Development Goal
SMS	- Short Message Service
UN	- United Nations
UNICEF	- United Nations Children’s Fund
UNSD	- United Nations Statistical Fund
WAPOR	- World Association for Public Opinion Research
WFP	- World Food Programme
WVS	- World Values Survey

Appendices

Appendix 1: List of SDG indicators targeting youth, that considered in this paper

*The indicators for which UNICEF is the sole or joint custodian of are highlighted in gray.

Target	Indicator	Proposed Tier
1.2 By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions	1.2.1 Proportion of population living below the national poverty line, by sex and age	Tier 1
	1.2.2 Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions	Tier 2
1.3 Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable.	1.3.1 Percentage of the population covered by social protection floors/systems disaggregated by sex, and distinguishing children, unemployed, old age, people with disabilities, pregnant women/newborns, work injury victims, poor and vulnerable	Tier 2
1.5 By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events and other economic, social and environmental shocks and disasters.	Number of deaths, missing people, injured, relocated or evacuated due to disasters per 100,000 people.	Tier 2
2.2 By 2030, end all forms of malnutrition, including achieving, by 2025, the internationally agreed targets on stunting and wasting in children under 5 years of age, and address the nutritional needs of adolescent girls, pregnant and lactating women and older persons	2.2.1 Prevalence of stunting (height for age <-2 SD from the median of the WHO Child Growth Standards) among children under five years of age	Tier 1
	2.2.2 Prevalence of malnutrition (weight for height >+2 or <-2 SD from the median of the WHO Child Growth Standards) among children under five, disaggregated by type (wasting and overweight)	Tier 1
3.1 By 2030, reduce the global maternal mortality ratio to less than 70 per 100,000 live births	3.1.1 Maternal deaths per 100,000 live births	Tier 2
	3.1.2 Proportion of births attended by skilled health personnel	Tier 1
3.2 By 2030, end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality to at least as low as 12 per 1,000 live births and under-5 mortality to at least as low as 25 per 1,000 live births	3.2.1 Under-five mortality rate (deaths per 1,000 live births)	Tier 1
	3.2.2 Neonatal mortality rate (deaths per 1,000 live births)	Tier 1
3.3 By 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne	3.3.1 Number of new HIV infections per 1,000 uninfected population (by age, sex, and key populations)	Tier 1

	3.3.3 Malaria incident cases per 1,000 persons per year	Tier 1
3.7 By 2030, ensure universal access to sexual and reproductive health-care services, including for family planning, information and education, and the integration of reproductive health into national strategies and programmes	3.7.1 Percentage of women of reproductive age (15-49 years) who have their need for family planning satisfied with modern methods	Tier 1
	3.7.2 Adolescent birth rate (10-14; 15-19) per 1,000 women in that age group	Tier 1
3.8 Achieve universal health coverage, including financial risk protection, access to quality essential health-care services and access to safe, effective, quality and affordable essential medicines and vaccines for all	3.8.1 Coverage of tracer interventions (e.g. child full immunization, ARV therapy, TB treatment, hypertension treatment, skilled attendant at birth, etc.)	Tier 3
3.9 By 2030, substantially reduce the number of deaths and illnesses from hazardous chemicals and air, water and soil pollution and contamination	3.9.1 Mortality rate attributed to household and ambient air pollution	Tier 1
4.1 By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes	4.1.1 Percentage of children/young people (i) in Grade 2/3, (ii) at the end of primary and (iii) at the end of lower secondary achieving at least a minimum proficiency level in (a) reading and (b) mathematics.	Tier 2
4.2 By 2030, ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education	4.2.1 Percentage of children under 5 years of age who are developmentally on track in health, learning and psychosocial well-being	Tier 2
	4.2.2 Participation rate in organised learning (one year before the official primary entry age)	Tier 1
4.5 By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training	4.5.1 Parity indices (female/male, urban/rural, bottom/top wealth quintile and others such as disability status and conflict-affected as data become available) for all indicators on this list that can be disaggregated	Tier 1/2/3 depending on indices
4.a Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all	4.a. Percentage of schools with access to (i) electricity; (ii) Internet for pedagogical purposes (iii) computers for pedagogical purposes; (iv) adapted infrastructure and material for students with disabilities; (v) basic drinking water and (vi) basic single-sex sanitation facilities; and (vii) basic handwashing facilities (as per the WASH indicator definitions)	Tier 1/2
5.2 Eliminate all forms of violence against all women and girls in the public and private spheres, including trafficking and sexual and other types of exploitation	5.2.1 Proportion of ever-partnered women and girls aged 15 years and over subjected to physical, sexual or psychological violence by a current or former intimate partner, in the last 12 months, by form of violence and by age group	Tier 2
	5.2.2 Proportion of women and girls aged 15 years and older subjected to sexual violence by persons other than an intimate partner, in the last 12 months, by age group and place of occurrence	Tier 2
5.3 Eliminate all harmful practices, such as child, early and forced marriage and female genital mutilation	5.3.1 Percentage of women aged 20-24 who were married or in a union before age 15 and before age 18	Tier 1

	5.3.2 Percentage of girls and women aged 15-49 years who have undergone FGM/C, by age group	Tier 1
6.1 By 2030, achieve universal and equitable access to safe and affordable drinking water for all	6.1.1 Percentage of population using safely managed drinking water services	Tier 1
6.2 By 2030, achieve access to adequate and equitable sanitation and hygiene for all and end open defecation, paying special attention to the needs of women and girls and those in vulnerable situations	6.2.1 Percentage of population using safely managed sanitation services including a hand washing facility with soap and water	Tier 1
7.1 By 2030, ensure universal access to affordable, reliable and modern energy services	7.1.2 Percentage of population with primary reliance on clean fuels and technology	Tier 1
8.7 Take immediate and effective measures to eradicate forced labour, end modern slavery and human trafficking and secure the prohibition and elimination of the worst forms of child labour, including recruitment and use of child soldiers, and by 2025 end child labour in all its forms	8.7.1 Percentage and number of children aged 5-17 years engaged in child labour, by sex and age group	Tier 1
16.1 Significantly reduce all forms of violence and related death rates everywhere	16.1.2 Conflict-related deaths per 100,000 population (disaggregated by age group, sex and cause)	Tier 2/Tier 3
16.2 End abuse, exploitation, trafficking and all forms of violence against and torture of children	16.2.1 Percentage of children aged 1-17 years who experienced any physical punishment and/or psychological aggression by caregivers in the past month	Tier 1
	16.2.3 Percentage of young women and men aged 18-24 years who experienced sexual violence by age 18	Tier 2
16.9 By 2030, provide legal identity for all, including birth registration	16.9.1 Percentage of children under 5 whose births have been registered with a civil authority, disaggregated by age	Tier 1

Appendix 2: Analytical Framework used to evaluate Tools

The table below lists the categories for each standard, on which the MTs were evaluated. Each category is broken down into descriptive questions to ensure comparable results after the evaluation.

Standard	Descriptive Indicators
Representation	● Aggregation: ○ Can respondent data be disaggregated to determine user profile? ● Sampling Frame: ○ What is the sampling frame? ● Sampling Design: ○ What is the sampling design?
Timeliness	● Collection Frequency: ○ How often are the perceptions gathered? ● Comparability: ○ For ongoing measurement, is the same population group surveyed? ○ For ongoing measurement, are the same questions asked? ● Data Processing Time: ○ What is the time lag between gathering data and processing results?
Measurement	● Lack of Recall Bias: ○ How many questions evaluate past feelings? ○ How many questions compare past perceptions with present? ● Lack of Social Desirability Bias: ○ How many questions ask for an opinion on the working of the government? ○ How many questions ask for an opinion on an activity considered illegal in the country? ● Lack of Framing effects: ○ How many questions are leading? ○ How many questions clearly state only one perception object? ● Lack of Proxy Effects: ○ How many questions ask about other people's behaviour? ● Strength of Response: ○ How many questions are closed? ○ How many questions utilize labelled scales to collect responses?
Credibility	● Translation: ○ Are questionnaires available in local languages?

	● Country Context: ○ For cross-national surveys, are country-specific questions included? ● Accessibility of Results: ○ How accessible are the results for key decision makers?

Appendix 3: Summary of Tools evaluated using Analytical Framework

This appendix provides an overview of all the MTs considered in this report along with their key characteristics.

Name	**Code**	**Ownership type**	**Focus SDGs**	**Questions considered**	**Data period considered**	**Mode of Collection**
Gallup World Poll	GSP	Private	3.3, 4, 5, 16.1	72/326	2014-16	FTF + Telephone
World Values Survey	WVS	NGO	5	28/250	2010-14	FTF
Global Attitudes Survey	GAS	NGO	1	61/731	2014-15	FTF + Telephone
European Social Survey	ESS	NGO	n.a	11/207	2014	FTF (CAPI)
Latino Barometer	LatBaro	NGO		28/219	2015	Tele/CAPI
Mobile Vulnerability Analysis and Monitoring	m-VAM	IGO	2	1/13		SMS
U-Report Nigeria	U-Rep(N)	IGO	2.2, 3, 4, 6,2	198	2014-17	SMS
U-Report Uganda	U-Rep(U)	IGO	3, 4, 5, 16	142/411	2014-17	SMS
U-Report Myanmar	U-Rep(M)	IGO		41/123	2016-17	Facebook Messenger

Appendix 4: Applying the Analytical Framework to the MTs

The following levels of symbols were used to denote the extent to which the MT demonstrated a particular standard. Numeric scores or rankings were not assigned to these MTs for a few reasons. Firstly the tools were not expressly designed for the purpose of measuring perceptions. Secondly the data was collected by the tools before the SDGs were fully formulated and none of the tools were expressly designed for the purpose of measuring progress to SDGs. Thirdly, the context within which each of the tools operate is varied with different purposes for their existence: from WVS which aims to provide scientific data on people's values to U-Report which is designed as an information dissemination and citizen engagement platform. Hence, to give the MTs rankings on measuring Perceptions for SDG monitoring would be an irrelevant exercise promising limited comparability.

✓✓✓ : Meets the ideal standard for the criteria

✓✓ : Has a few minor operational flaws

✓ : Most aspects of the criteria are unfulfilled

<blank>: Data is unavailable or information is completely absent

Appendix 5: The Representation-Timeliness Trade-off

	GWP	WVS	GAS	ESS	Lat Baro	U-Rep (N)	U-Rep (U)	U-Rep (M)
Aggregation	✓✓	✓✓ ✓	✓✓	✓✓ ✓	✓✓✓	✓✓	✓✓	✓✓
Sampling Frame	✓✓	✓✓	✓✓	✓✓ ✓	✓✓	✓	✓	✓
Sampling Design	✓✓	✓✓	✓✓	✓✓ ✓	✓✓	✓	✓	✓
Collection Frequency	✓	✓	✓	✓	✓	✓✓✓	✓✓✓	✓✓✓
Comparability	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓
Data processing time	✓	✓	✓	✓	✓	✓✓✓	✓✓✓	✓✓✓

Appendix 6: Evaluation of Tools for the Measurement Standard

	GWP	WVS	GAS	ESS	LatBaro	U-Rep (N)	U-Rep (U)	U-Rep (M)
Lack of Recall bias	✓✓ ✓	✓✓ ✓	✓✓ ✓	✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓
Lack of Social Desirability Bias	✓✓	✓✓	✓	✓✓ ✓	✓	✓✓	✓	✓✓
Lack of Framing Effects	✓✓	✓✓ ✓	✓✓ ✓	✓✓ ✓	✓✓	✓✓	✓	✓✓
Lack of Proxy Effects	✓✓ ✓	✓✓ ✓	✓✓ ✓	✓✓ ✓	✓✓✓	✓✓✓	✓✓	✓✓✓
Strength of Response	✓✓	✓✓ ✓	✓✓	✓✓ ✓	✓✓	✓✓	✓	✓

Appendix 7: Evaluating the Credibility of the Tools

	GWP	WVS	GAS	ESS	LatBaro	U-Rep (N)	U-Rep (U)	U-Rep (M)
Translation	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓
Country context	✓✓	✓✓	✓✓	✓	✓✓	✓✓✓	✓✓✓	✓✓✓
Accessibility of Results	✓	✓✓	✓✓	✓✓	✓✓	✓✓✓	✓✓✓	✓✓✓

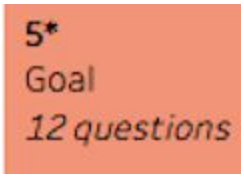
Appendix 8: Mapping Tools to SDGs

I = Indicators T= Targets G=Goals

	WVS			GWP			GAS			ESS			LatinoBarometer			U-Report Nigeria			U-Report Uganda			U-Report Myanmar		
	I	T	G	I	T	G	I	T	G	I	T	G	I	T	G	I	T	G	I	T	G	I	T	G
1.2.1																2								
1.2.2					6						2			3		3	2							
1.3.1	1										2													
1.5.1			1						24						8					4	14		4	7
2.2.1				2														1						
2.2.2			1	2					5								18			2	2			
3.1.1																								
3.1.2										1							4			5				
3.2.1										1														
3.2.2					7											1	16			4				
3.3.1				6																		2		
3.3.3					8						1					1	18			9				
3.7.1				1			1																	
3.7.2					1			1						1			6			9			2	
3.9.1			3	1	3	3	3	2	6		1	1			2			34			29			6
4.1.1				1	7											8	1							
4.2.1				1	7											2	5			6				
4.5.1					1			1									8			1				
4.a.1			2		4	8		2	5			1			7			17		5	41		2	8
5.2.1	1			1												2						1		
5.2.2					4					1	1					2	5			17			3	
5.3.1				1												1			3			1		
5.3.2			12			8			3			2			6	3	3	4	12		28			2
6.1.1				3	2			1						1		3	3		1					
6.2.1			1					2						1		1	18				1	1		1
7.1.2								1						2							4			
8.7.1			3	1																	9			
16.1.2		5			9												12			2			2	
16.2.1	1															1						1		
16.2.3																4	9			12		6	1	
16.9.1			2											1	5	1					19			5

Appendix 8.a: Diversity of SDG-related perceptions questions asked

Area is proportional to the percentage of perception questions related to each goal, target or indicator (see above chart for specific numerical values). Colors are variants of official SDG colors.

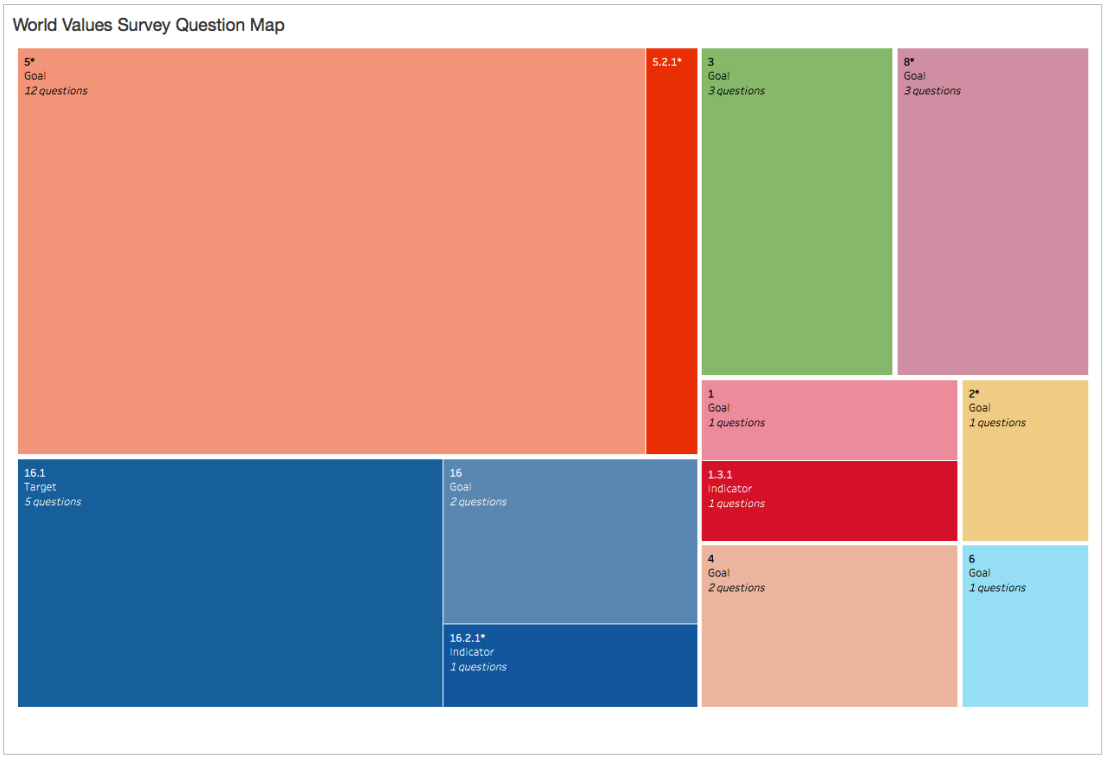


Which Goal, Target, or Indicator (GTI)

Level

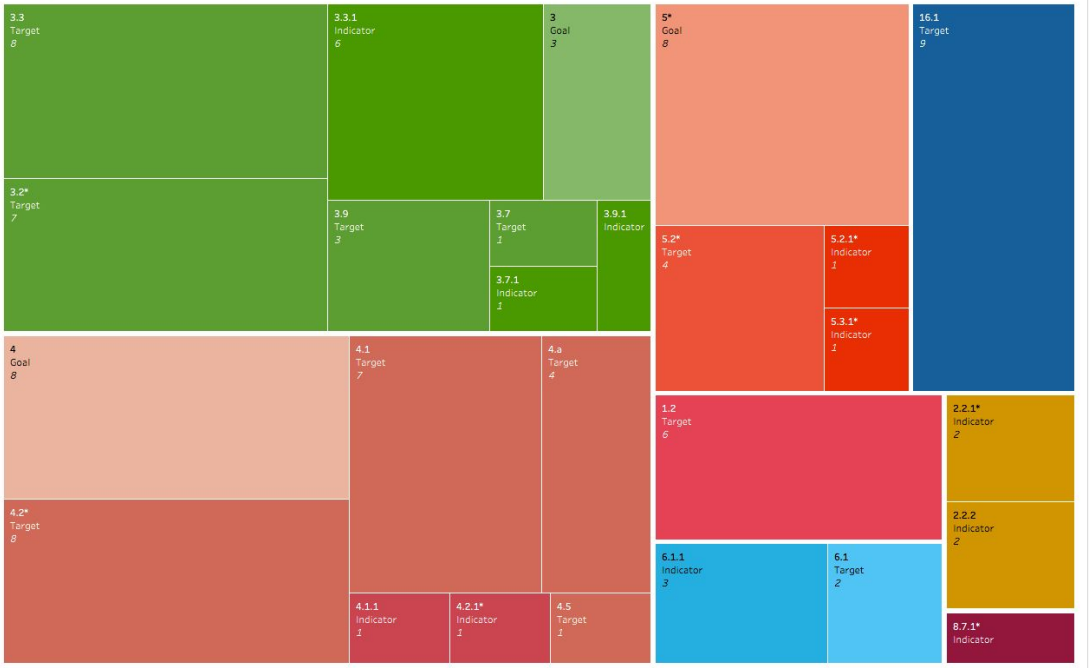
Number of perception questions related to this GTI

* indicates that Unicef is sole or joint custodian for this GTI



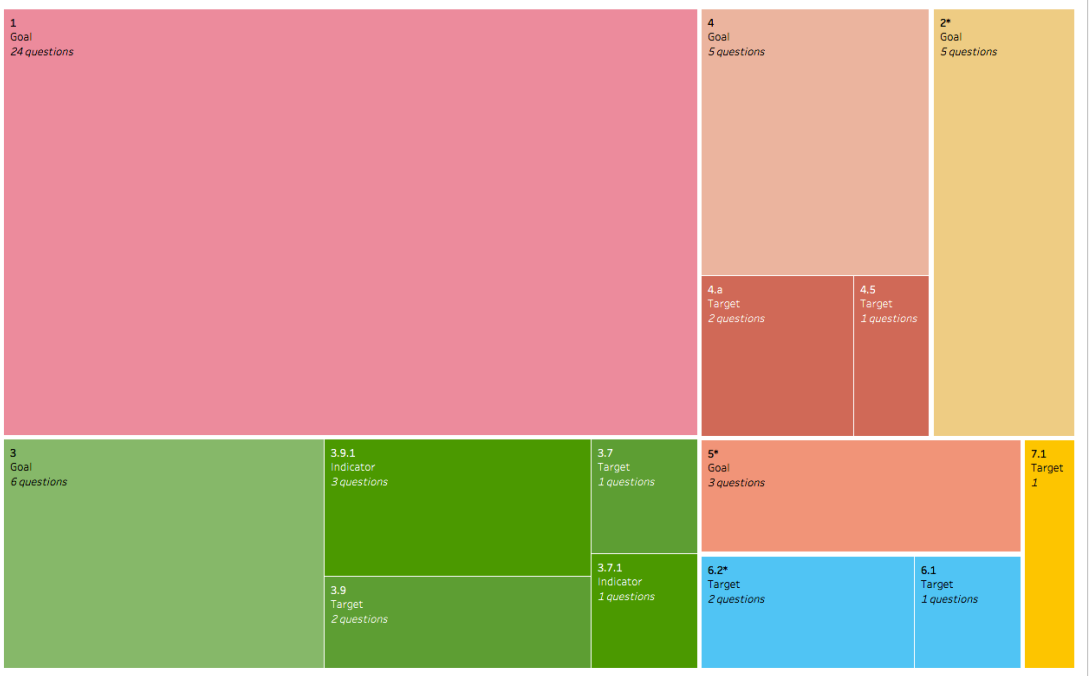
Color Key

Gallup World Poll Question Map

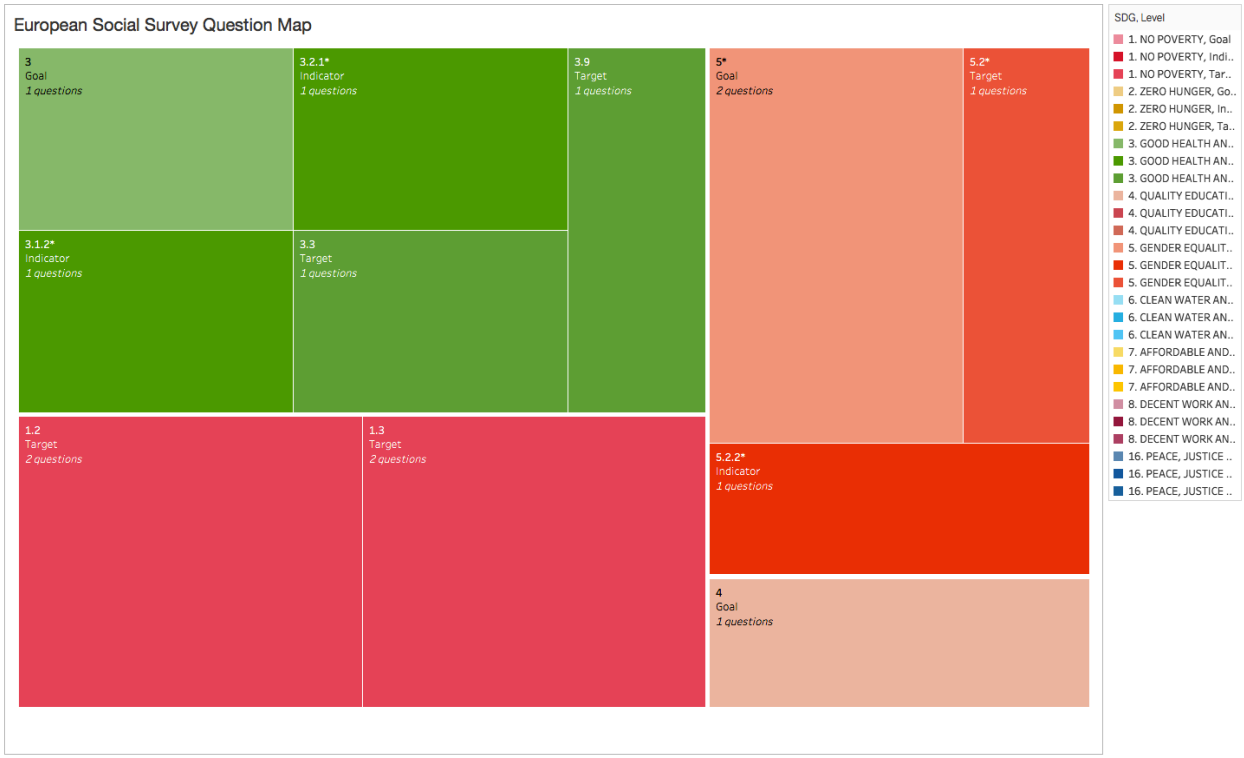


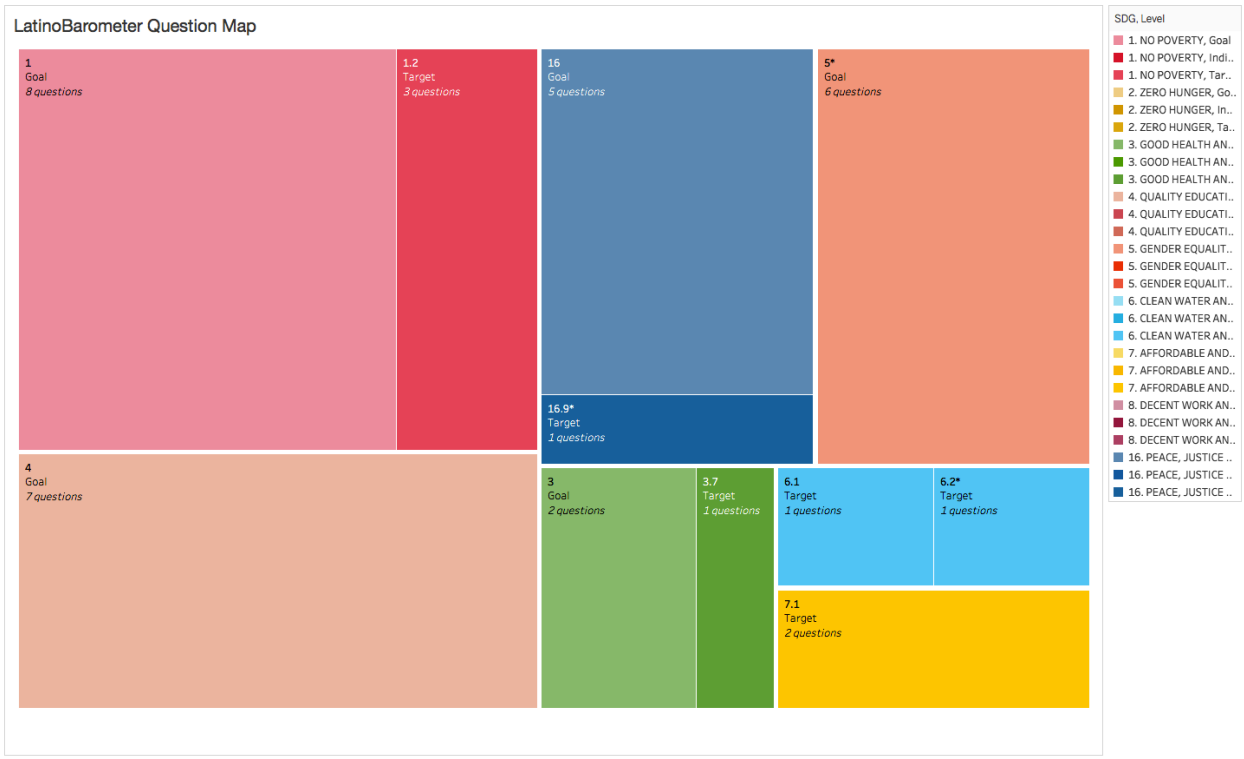
SDG, level
- 1. NO POVERTY, Goal
- 1. NO POVERTY, Indi..
- 1. NO POVERTY, Tar..
- 2. ZERO HUNGER, Go..
- 2. ZERO HUNGER, In..
- 2. ZERO HUNGER, Ta..
- 3. GOOD HEALTH AN..
- 3. GOOD HEALTH AN..
- 3. GOOD HEALTH AN..
- 4. QUALITY EDUCATI..
- 4. QUALITY EDUCATI..
- 4. QUALITY EDUCATI..
- 5. GENDER EQUALIT..
- 5. GENDER EQUALIT..
- 5. GENDER EQUALIT..
- 6. CLEAN WATER AN..
- 6. CLEAN WATER AN..
- 6. CLEAN WATER AN..
- 7. AFFORDABLE AND..
- 7. AFFORDABLE AND..
- 7. AFFORDABLE AND..
- 8. DECENT WORK AN..
- 8. DECENT WORK AN..
- 8. DECENT WORK AN..
- 16. PEACE, JUSTICE ..
- 16. PEACE, JUSTICE ..
- 16. PEACE, JUSTICE ..

Global Attitude Survey Question Map

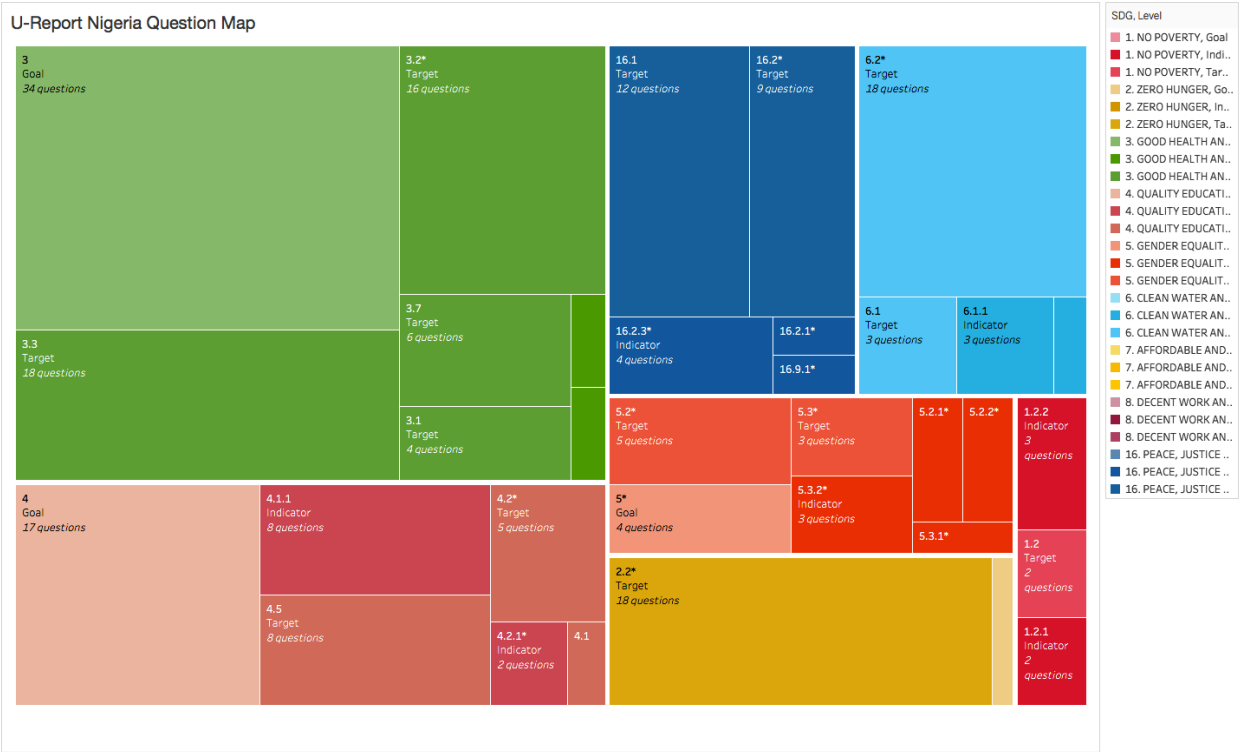


SDG, Level
- 1. NO POVERTY, Goal
- 1. NO POVERTY, Indi..
- 1. NO POVERTY, Tar..
- 2. ZERO HUNGER, Go..
- 2. ZERO HUNGER, In..
- 2. ZERO HUNGER, Ta..
- 3. GOOD HEALTH AN..
- 3. GOOD HEALTH AN..
- 3. GOOD HEALTH AN..
- 4. QUALITY EDUCATI..
- 4. QUALITY EDUCATI..
- 4. QUALITY EDUCATI..
- 5. GENDER EQUALIT..
- 5. GENDER EQUALIT..
- 5. GENDER EQUALIT..
- 6. CLEAN WATER AN..
- 6. CLEAN WATER AN..
- 6. CLEAN WATER AN..
- 7. AFFORDABLE AND..
- 7. AFFORDABLE AND..
- 7. AFFORDABLE AND..
- 8. DECENT WORK AN..
- 8. DECENT WORK AN..
- 8. DECENT WORK AN..
- 16. PEACE, JUSTICE ..
- 16. PEACE, JUSTICE ..
- 16. PEACE, JUSTICE ..

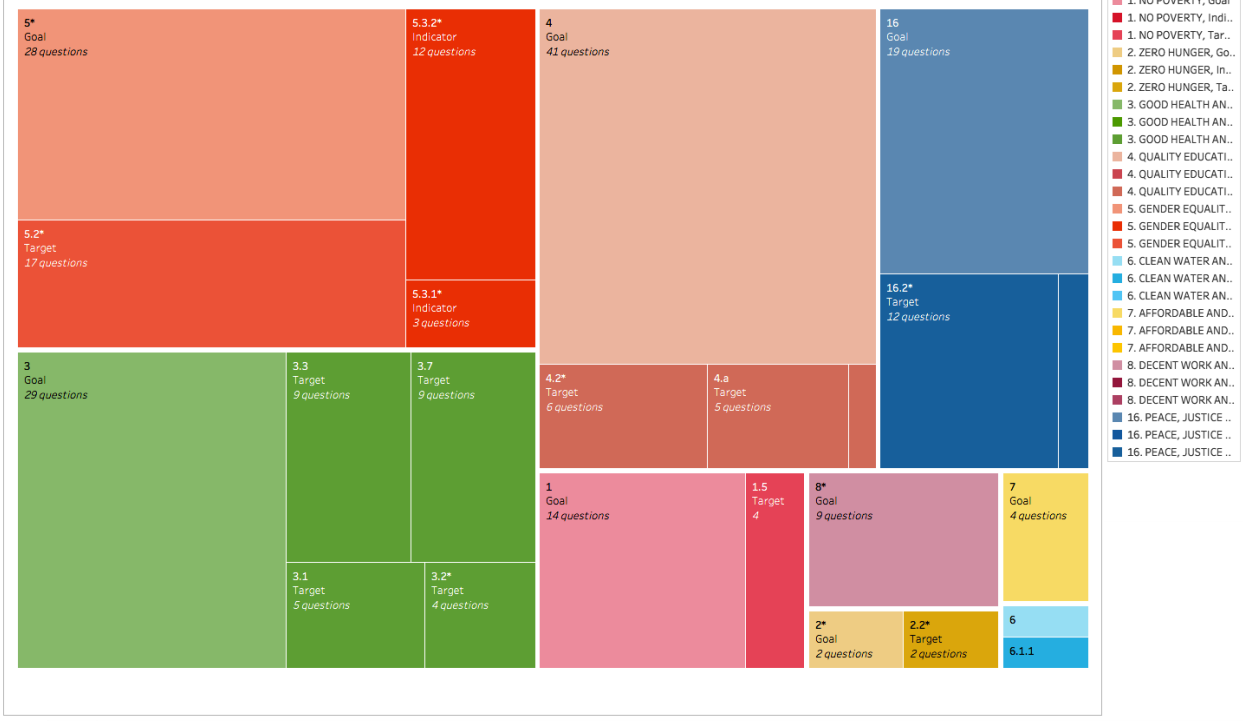

European Social Survey Question Map
3
Goal
1 questions
3.2.1*
Indicator
1 questions
3.9
Target
1 questions
5*
Goal
2 questions
5.2*
Target
1 questions
3.1.2*
Indicator
1 questions
3.3
Target
1 questions
1.2
Target
2 questions
1.3
Target
2 questions
5.2.2*
Indicator
1 questions
4
Goal
1 questions
SDG, Level
1. NO POVERTY, Goal
1. NO POVERTY, Indi..
1. NO POVERTY, Tar..
2. ZERO HUNGER, Go..
2. ZERO HUNGER, In..
2. ZERO HUNGER, Ta..
3. GOOD HEALTH AN..
3. GOOD HEALTH AN..
3. GOOD HEALTH AN..
4. QUALITY EDUCATI..
4. QUALITY EDUCATI..
4. QUALITY EDUCATI..
5. GENDER EQUALIT..
5. GENDER EQUALIT..
5. GENDER EQUALIT..
6. CLEAN WATER AN..
6. CLEAN WATER AN..
6. CLEAN WATER AN..
7. AFFORDABLE AND..
7. AFFORDABLE AND..
7. AFFORDABLE AND..
8. DECENT WORK AN..
8. DECENT WORK AN..
8. DECENT WORK AN..
16. PEACE, JUSTICE ..
16. PEACE, JUSTICE ..
16. PEACE, JUSTICE ..

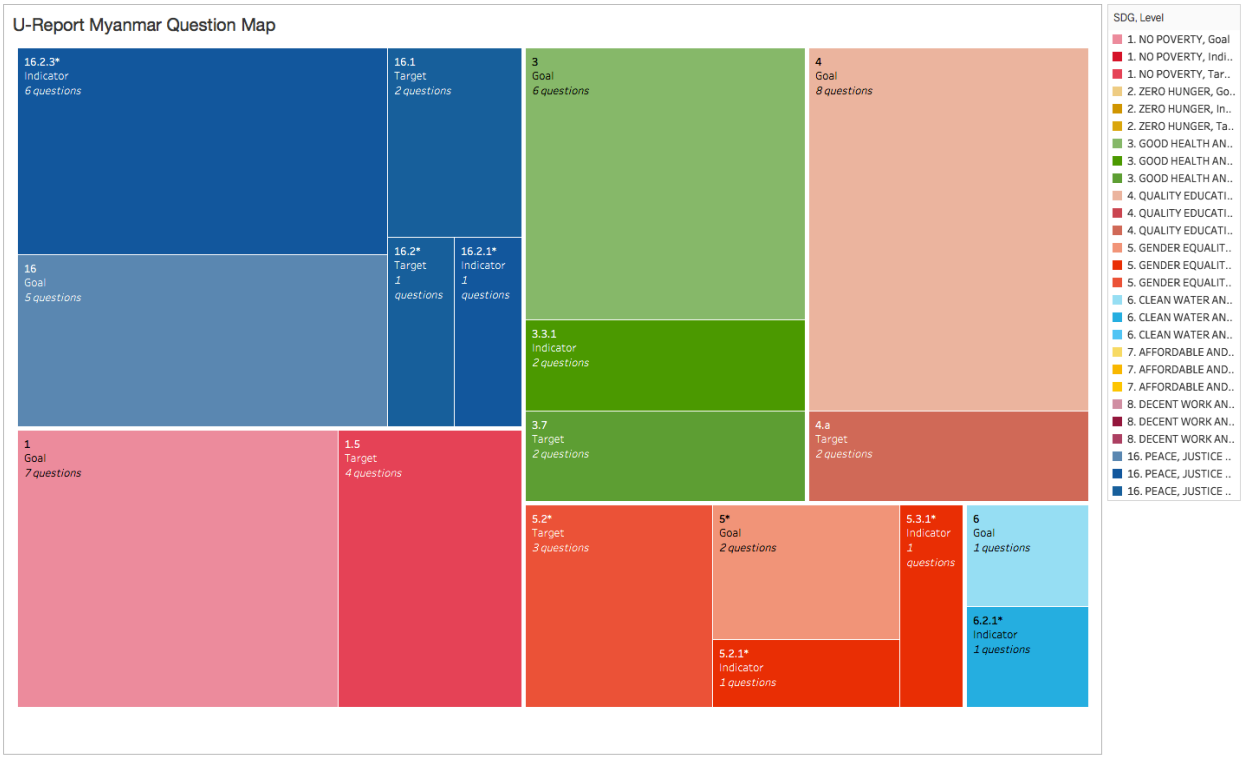

LatinoBarometer Question Map
1
Goal
8 questions
1.2
Target
3 questions
16
Goal
5 questions
5*
Goal
6 questions
16.9*
Target
1 questions
4
Goal
7 questions
3
Goal
2 questions
3.7
Target
1 questions
6.1
Target
1 questions
6.2*
Target
1 questions
7.1
Target
2 questions
SDG, Level
1. NO POVERTY, Goal
1. NO POVERTY, Indi..
1. NO POVERTY, Tar..
2. ZERO HUNGER, Go..
2. ZERO HUNGER, In..
2. ZERO HUNGER, Ta..
3. GOOD HEALTH AN..
3. GOOD HEALTH AN..
3. GOOD HEALTH AN..
4. QUALITY EDUCATI..
4. QUALITY EDUCATI..
4. QUALITY EDUCATI..
5. GENDER EQUALIT..
5. GENDER EQUALIT..
5. GENDER EQUALIT..
6. CLEAN WATER AN..
6. CLEAN WATER AN..
6. CLEAN WATER AN..
7. AFFORDABLE AND..
7. AFFORDABLE AND..
7. AFFORDABLE AND..
8. DECENT WORK AN..
8. DECENT WORK AN..
8. DECENT WORK AN..
16. PEACE, JUSTICE ..
16. PEACE, JUSTICE ..
16. PEACE, JUSTICE ..

U-Report Nigeria Question Map



U-Report Uganda Question Map

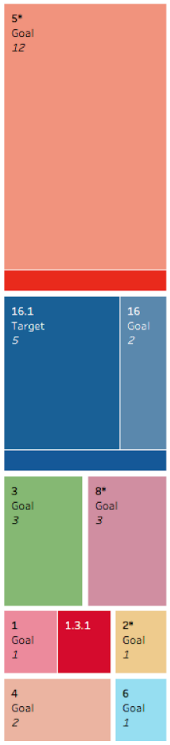


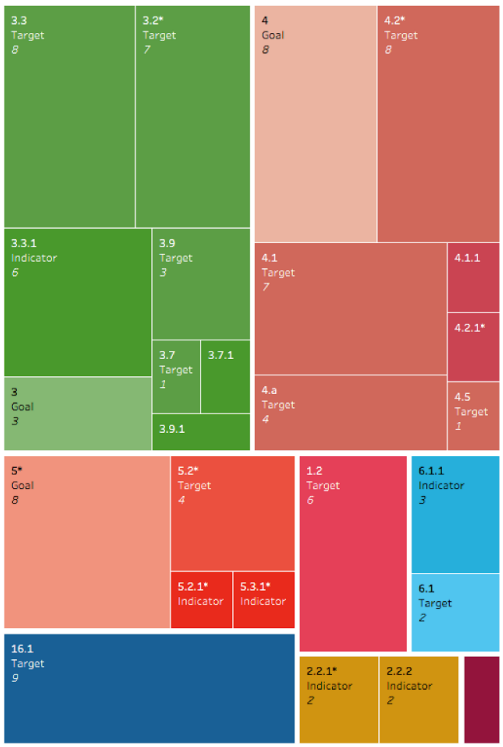


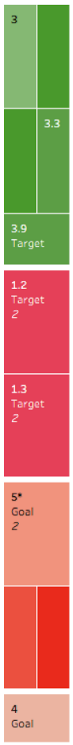
Appendix 8.b: Quantitative comparison of tools

Areas are proportional to the amount of SDG-related perceptions questions asked since 2014.

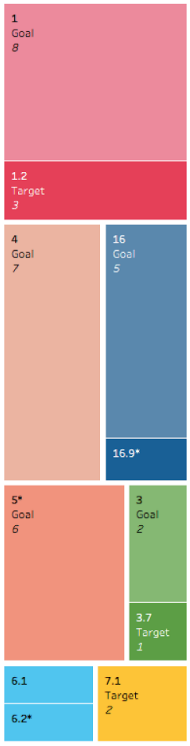
WVS (33 questions) GWP (99) ESS (14)



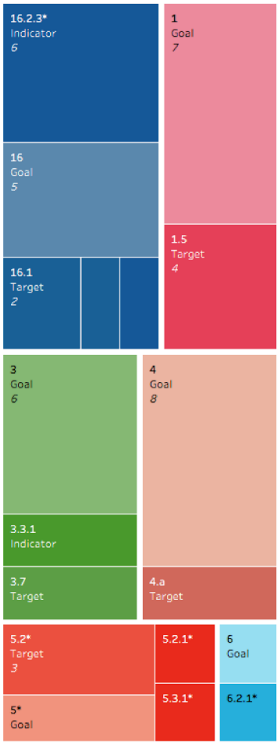




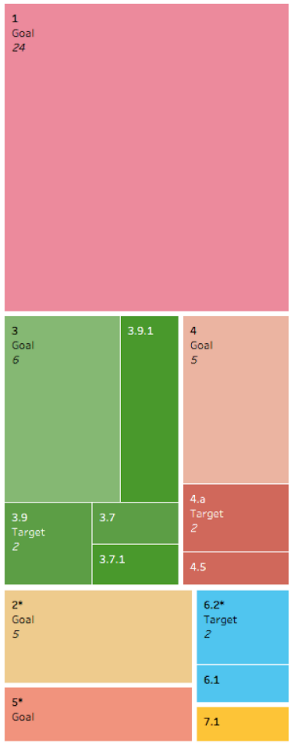
LatinoBarometer (37)



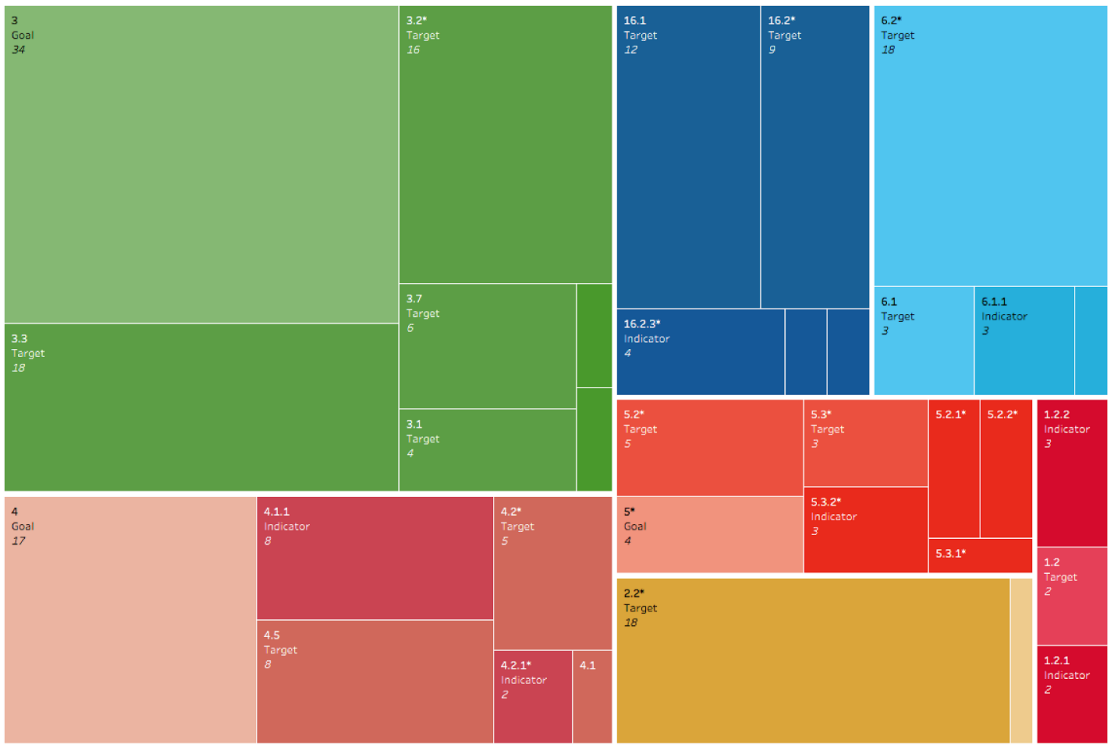
U-Report Myanmar (55)



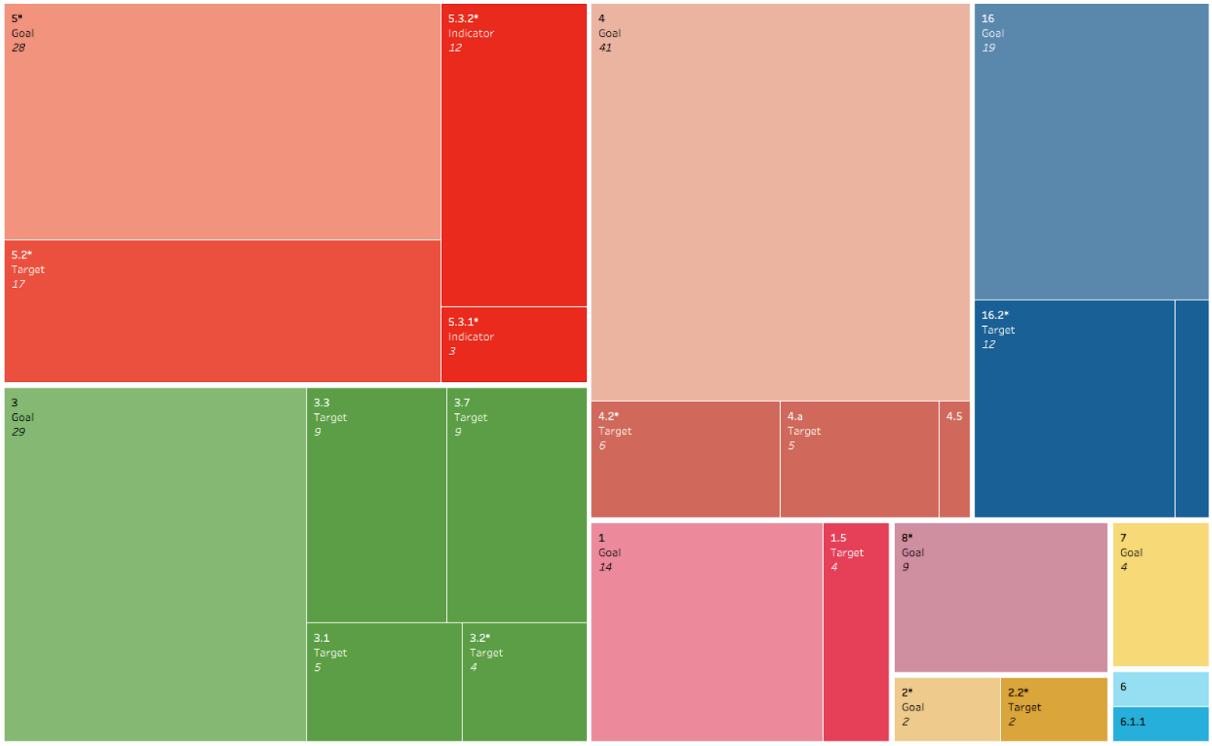
GAS (57)



U-Report Nigeria (219)



U-Report Uganda (239)



Acknowledgements

The research team is indebted to a range of experts, practitioners, and academics who generously donated their time to this paper. In particular the team would like James Powell at UNICEF for his passion and dedication as our research partner. We would also like to thank Laura Rodriguez Takeuchi at the University of Manchester for her detailed feedback on the proposed Analytical Framework and paper, and Professor Vu Minh Khuong from the Lee Kuan Yew School of Public Policy for his regular advice as our faculty advisor. Furthermore, we're indebted to Alexandra Silfverstolpe (Data Act Labs), Cary McCormick (UNICEF), Davis Adieno (CIVICUS), Francelino Murela (UNICEF), Jean-Martin Bauer (WFP), Jonathan Glennie (IPSOS), Klaus Oberbauer (UNICEF), Namrata Chindarkar (LKYSPP), Riccardo Polastro (UNICEF), Sandar Linn (UNICEF), Sébastien Hine (Save the Children), Sharad Sapra (UNICEF), Victoria Ndoh (UNICEF) and Zeger van der Wal (LKYSPP) for generously giving their time to be interviewed and providing valuable constructive feedback on our paper, and their insights on our research question. Finally, thank you to Kseniya Kizilova from the World Values Survey Association Secretariat for providing information about the forthcoming wave of the WVS and how their approach towards the SDGs, and to Alissa Fasman (LKYSPP) for providing the original introduction to the UNICEF Global Innovation Centre, and enabling this project to get started.

Bibliography

Alkire, S., & Samman, E. (2014). *Mobilising the household data required to progress toward the SDGs* (No. ophiwp072). Queen Elizabeth House, University of Oxford.

United Nations General Assembly (2015). Transforming our world: the 2030 Agenda for Sustainable Development. *New York: United Nations*.

Bamberger, Michael, Linda Raftree, and Veronica Olazabal. "The role of new information and communication technologies in equity-focused evaluation: Opportunities and challenges." *Evaluation* 22, no. 2 (2016): 228-244.

Barrett, C. B. (2010). Measuring food insecurity. *Science*, *327* (5967), 825-828.

Benjamin Leo. 2013. “Is Anyone Listening? Does US Foreign Assistances Target People’s Top Priorities?.” CGD Working Paper 348. Washington, DC: Center for Global Development.

Chai, J., & Cummins, M. (2014). From 'What Happened?' to 'What's Happening.”

Dabalen, Andrew, Alvin Etang, Johannes Hoogeveen, Elvis Mushi, Youdi Schipper, and Johannes von Engelhardt. "Mobile Phone Panel Surveys in Developing Countries." (2016).

DeLeon, P. (2006). The Historical Roots of the Field. In M. Moran, M. Rein, & R. E. Goodin (Eds.), *The Oxford Handbook of Public Policy*. Oxford University Press.

Diener, E., Oishi, S., & Lucas, R. E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual review of psychology*, *54*(1), 403-425.

ESOMAR/WAPOR (2014). “ESOMAR / WAPOR Guidelines on Opinion Polls and Published Surveys.”

Fischer, F. (2003). Making Social Science Relevant: Policy Inquiry in Critical Perspective. In *Reframing Public Policy: Discursive Politics and Deliberative Practices*. Oxford University Press.

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology* (Vol. 561). John Wiley & Sons.

Herbert, S. (2013). Perception Surveys in Fragile and Conflict-affected States. *GSRDC Applied Knowledge Services.*

How SMS helps prevent HIV in Mozambique. (2016, August 3). Retrieved October 24, 2017, from http://unicefstories.org/2016/08/03/how-sms-helps-prevent-hiv-in-mozambique

Howard, J., & Wheeler, J. (2015). What community development and citizen participation should contribute to the new global framework for sustainable development. *Community Development Journal*, *50*(4), 552-570.

Independent Expert Advisory Group on a Data Revolution for Sustainable Development. (2014). A world that counts: mobilizing the data revolution for sustainable development.

IAEG-SDGs. (n.d.). Retrieved November 2, 2016, from https://unstats.un.org/sdgs/iaeg-sdgs/

Leo, B., & Morello, R. (2015). Asking What the People Want: Using Mobile Phone Surveys to Identify Citizen Priorities.

Lindblom, C. E., & Cohen, D. K. (1979). *Usable Knowledge: Social Science and Social Problem Solving*. Yale University Press.

Ling, T., & Villalba van Dijk, L. (2009). Performance Audit Handbook.

Link, M. W., Murphy, J., Schober, M. F., Buskirk, T. D., Childs, J. H., & Tesfaye, C. L. (2014). Mobile technologies for conducting, augmenting and potentially replacing surveys executive summary of the AAPOR task force on emerging technologies in public opinion research. *Public Opinion Quarterly*, nfu054.

Lüge, Timo (2015). “Comparison of SMS Platforms.” UNHCR Public Health Commission.

Morrow, Nathan, Nancy Mock, Jean-Martin Bauer, and Jennifer Browning. "Knowing Just in Time: Use cases for mobile surveys in the humanitarian world." *Procedia Engineering* 159 (2016): 210-216.

OECD Better Life Initiative. (2013). OECD guidelines on measuring subjective well-being.

Peixoto, T., & Fox, J. (2016). When Does ICT-Enabled Citizen Voice Lead to Government Responsiveness?. *IDS Bulletin*, *47*(1).

Price, V. (2007). The Public and Public Opinion in Political Theories. In Donsbach, W. & Traugott, M.W. (1st ed.). The SAGE Handbook of Public Opinion Research. London: SAGE.

Prpić, J., Taeihagh, A. and Melton, J. (2015). "The Fundamentals of Policy Crowdsourcing." *Policy & Internet, 7*(3).

Revised list of global Sustainable Development Goal indicators. (2017, March). Retrieved March 30, 2017, from https://unstats.un.org/sdgs/indicators/Official%20Revised%20List%20of%20global%20SDG%20indicators.pdf

Rodriguez Takeuchi, Laura, Hine, Sebastien, and Chavez, Cirenia. 2015. "Working Paper 413: Asking People What They Think: Using Perceptions Data to Monitor the Post-2015 Agenda." Overseas Development Institute.

Seaford, C. (2013). The multiple uses of subjective well-being indicators. *Social indicators research*, *114*(1), 29-43.

Sharma, Y. (2014, December 4). Q&A: The best ways to track the SDGs. Retrieved October 18, 2016, from http://www.scidev.net/global/mdgs/feature/track-sdgs-data-development-goals.html

Shaw, A., Nguyen, L., Nischan, U., & Sy, H. (2011). Comparative Assessment of Software Programs for the Development of Computer-Assisted Personal Interview (CAPI) Applications. Technical report, The World Bank Living Standards and Measurement Study.

Stiglitz, J., Sen, A., & Fitoussi, J.-P. (2009). The measurement of economic performance and social progress revisited. Reflections and overview. *Commission on the Measurement of Economic Performance and Social Progress, Paris*.

World Food Programme (n.d.). "Introduction to mVAM". VAM Resource Centre. Accessed September 11, 2016. <http://resources.vam.wfp.org/mVAM>

"U-Report." U-Report. Accessed September 11, 2016. https://community.rapidpro.io/stories/u-report/.

Vandemoortele, Jan (2011) If not the Millennium Development Goals, then what?, *Third World Quarterly*, *32*(1), 9-25

Veenhoven, R. (2015). Social conditions for human happiness: A review of research. *International Journal of Psychology*, *50*(5), 379-391.

Weiss, C. H. (1979). The Many Meanings of Research Utilization. *Public Administration Review*, *39*(5), 426–431.