Working paper 413

Asking people what they think

Using perceptions data to monitor the post-2015 agenda

Laura Rodríguez Takeuchi and Sébastien Hine with Cirenia Chávez

Key messages

- The call for a 'data revolution' has spurred debate around the inclusion of new data and indicators to measure progress towards development goals. Indicators of perceptions – based on asking people what matters to them most and their opinions of change – could help to stimulate public debate and hold policy-makers accountable.
- Key strengths of perceptions data are their timeliness and frequency – such attributes could make them very useful as warning signals for policy intervention.
- We illustrate the potential of perception indicators in three post-2015 areas: social norms related to gender, violence and security, and governance. Perceptions and so-called 'objective' data can measure complementary aspects of these areas. Analysing gaps between perceptions and objective indicators can improve understanding of how people are dissatisfied, or when there are implementation gaps in the policies intended to tackle these areas.
- Main limitations of this data are the challenge of ensuring the reliability of the information obtained, and difficulties in making meaningful comparisons across groups of people. We suggest that perceptions data would be more useful to monitor changing situations over time within countries, rather than to establish comparisons across them.

Contents

List of tables, figures and boxes

Tables

Figures

Boxes

Acronyms and abbreviations

DHS	Demographic and Health Surveys
FAO	Food and Agriculture Organization
FGM	female genital mutilation
GWP	Gallup World Poll
HDI	Human Development Index
HLP	High Level Panel of Eminent Persons on the Post-2015 Development Agenda (UN)
IAEG-GS	Inter-Agency and Expert Group on Gender Statistics (UN)
IEAG	Independent Expert Advisory Group on a Data Revolution for Sustainable Development
IPFRI	International Food Policy Research Institute
LSMS	Living Standard Measurement Survey
MDG	Millennium Development Goals
MICS	Multiple Indicator Cluster Surveys
NSO	National Statistical Office
OECD	Organisation for Economic Co-operation and Development
OPHI	Oxford Poverty & Human Development Initiative
OWG	Open Working Group
SDG	Sustainable Development Goals
SIGI	Social Institutions and Gender Index (OECD)
SWB	subjective wellbeing
UNDP	United Nations Development Programme
WEIA	Women's Empowerment in Agriculture Index (IPFRI)
WVS	World Values Surveys

Acknowledgements

We are grateful for helpful comments and inputs from Emma Samman, Jessica Hagen-Zanker and Claire Melamed from ODI as well as from Laura Camfield (UEA), Hayk Gyuzalyan (Ipsos Mori) and Enrico Giovanini (University of Rome 'Tor Vergata'). Nina Behrman provided editorial support. The usual disclaimers apply.

Introduction and context

The process of developing and assessing progress against the Millennium Development Goals (MDG) has been criticised as technocratic and too much 'behind closed doors'. In contrast, negotiations around the successors to the MDG have been wide-ranging, seeking to be consultative and incorporating the views of all governments and of citizens globally. This spirit of making the post-2015 process more inclusive and responsive to citizen's views and demands should be carried into the monitoring of the new development goals.

Perceptions data are increasingly widely available, and could be useful in monitoring progress towards new development goals. Available data include those from:

- the World Value Surveys (WVS) conducted in over 80 countries and the Gallup World Poll (GWP) covering over 160 countries
- the 'barometers' (Latin America, Asia, Arab region and Africa), the Latin America Public Opinion Project and the European Social Survey, which have a more regional focus
- questions in Living Standard Measurement Surveys (LSMS), which are internationally comparable multi-topic household surveys conducted in developing countries
- subjective wellbeing data (SWB) collected by National Statistical Offices (NSOs) in some countries[1].

While existing data could be used as criteria for narrowing down the plethora of proposed goals and targets, the vast interest in examining new forms of measuring – what to measure, and how to measure it – provides a good opportunity to examine new possibilities. The UN High Level Panel of Eminent Persons on the Post-2015 Development Agenda (HLP) has called for a 'data revolution', and the recent report of the UN's Expert Group on the Data Revolution (IEAG, 2014) prompted wide debate on possible new approaches and incorporating new methods of collecting and using different types of data.

'[A data revolution is] an explosion in the volume of data, the speed with which data are produced, the number of producers of data, the dissemination of data, and the range of things on which there is data, coming from new technologies such as mobile phones and the "internet of things", and from other sources, such as qualitative data, citizen-generated data and perceptions data.' (IEAG, 2014)

The Secretary-General's synthesis report (UN, 2014b) also clearly underlines the importance of subjective data in measuring wellbeing. As argued by the Stiglitz, Sen, Fitoussi Commission (2009), perceptions data could shift the measurement debate towards improving wellbeing of citizens around the world as they understand it. Nevertheless, to our knowledge, there is no specific assessment of the potential of subjective data, or the challenges involved, in this context. This paper addresses this issue, framed in the context of the monitoring of the new set of development goals. It also outlines some methodological issues more broadly.

Section 2 defines perceptions data and presents general initial thoughts on their potential. Section 3 considers the potential of perceptions data for monitoring the evolution of gender norms, violence and security, and governance. Section 4 concentrates on the challenges, and in Section 5 we focus on practical considerations for the inclusion of perceptions data in the monitoring of post-2015 development goals.

1 National Statistical Offices including SWB measures are: Canada, France, Italy, Mexico, Morocco, New Zealand, UK, US, and most EU countries (O'Donnell et al., 2014).

The case for perceptions data

Perceptions data, often called subjective data, concern the expressed views (opinions or perceptions)[2] of people on a particular topic.[3] In contrast, so-called 'objective data' are reports of material possessions, events or situations collected through direct physical examination or observation. We refer to this second kind of data as 'objective' in this paper, although we recognise that they can also be subject to biases and different interpretations.

The definition we adopt also excludes self-reported objective data, or what survey researchers call answers to 'factual' questions. These refer to information on past actions or objective living conditions, for example voting behaviour or self-reported income, although both may be collected through the same means of survey questionnaire. This distinction is important. Take for example, a question asking an individual to report their household income (either as an exact figure or within a set range). This type of question is self-reported, rather than collected for example through tax records, but it is not about perceptions as defined here.

While clearly there are some objectively measurable things such as an amount of food (or the necessary income to purchase that food), that people need to lead full and healthy lives, people can also make subjective assessments of these objective conditions. The main advantage of perceptions data is that they provide 'the respondents' own views directly … perceptions of themselves and their world, which are unobtainable in any other way' (Baker et al., 2002). For example, a question about the perception of the household income *sufficiency* would be included in our definition of perceptions data (Box 1). These subjective judgements matter regardless of the objective circumstances because people may react to it if they can – with consequences for how they respond to incentives, and to political systems and processes, and for how people relate to each other (Coulthard et al., 2014). Moreover, there are other things that are best assessed only in subjective terms because what really matters is how people experience these (Veenhoven, 2004; Gough and McGregor, 2009; Coulthard et al., 2014). Veenhoven (2004) uses the example of xenophobia. To a degree, this can be measured by a reduction of racist attacks or the rates of interethnic marriage, for instance, but perceptions are needed to capture this concept fully. Similar arguments apply for the concepts of human dignity and security (Coulthard et al., 2014). Perception data are collected mainly through household surveys. This paper focuses on internationally comparable data that could be used for monitoring the agreed Sustainable Development Goals (SDG)/post-2015 targets. Some multi-topic household surveys include selected questions to ascertain perceptions around specific issues. Subjective modules can be appended to broad household surveys, as in the 'missing dimensions' project of the Oxford Poverty & Human Development Initiative (OPHI).[4] There are some specialised surveys focusing exclusively on capturing perceptions data. Two of the largest of these are the WVS and the GWP, conducted in a wide number of countries around the world following a standard methodology and questionnaire. They ask a representative sample of participating country residents about their perceptions on topics ranging from government and politics, to family, religion, ethics and wellbeing. The regional barometer surveys are similar but the specific topics and questions may vary in the regional surveys.

The UN MyWorld survey is another large-scale perception survey. Resembling more a short opinion poll than a household survey, it consists of only one core question ('*Which of these are most important for you and your family?*'[5]), aimed at gathering people's priorities for global development post-2015. In addition, although there

2 Even within the definition adopted, there can be various types of data. A distinction could be made between questions that ask about perceptions ('how do you feel about …?'), opinions ('do you think that …?') and choices ('should the government do ... or …?'). The main emphasis is on the first two of these types, especially since the third may have different methodological problems. That said, some issues are common to all perceptions data, and even some to all estimates based on survey methods.

3 Experts' valuations could be classified as perceptions data, insofar as they refer to opinions/perceptions. Although expert views have been widely used, for example in the measurement of governance, we exclude them from the definition used in this paper as our main interest here is perceptions data drawn from (mostly representative) samples of the general population rather than from selected groups of experts.

4 http://www.ophi.org.uk/research/missing-dimensions/

5 The survey asks respondents to choose 6 issues that matter most out of 16 possibilities.

Box 1: Income perception questions

Household or personal income is often used as a measure of welfare and subsequent estimates of monetary poverty and inequality. It is also possible to derive poverty and inequality measures from subjective perceptions of relative wealth/poverty rather than actual income. Some household surveys have incorporated this type of question, for example:

- How does your household income compare with other households in your village/neighbourhood?
  Much above average income
  Above average income
  Average income
  Below average income
  Much below average income
  (Do not know)
- What monthly income level do you consider to be minimal for your household, i.e. your household could not make ends meet with less?
- Is the total monthly income of your household higher, lower or more or less the same as this figure?
  Much higher
  Higher
  More or less the same
  Lower
  Much lower
  (South African Social Attitudes Survey, 2007)
- Your household income:
  Is not enough to meet the minimum expenditure
  Is just enough to meet the minimum expenditure
  Is above the minimum expenditure
  Don't know/NA
  (Encuesta de Calidad de Vida de Bogotá, 2007)
- Is your pay just? We are not asking about how much you would like to earn – but what you feel is just given your skills and effort. If you are not working now, please tell about your last job.
  Much less than is just
  A little less than is just
  About just for me
  A little more than is just
  Much more than is just
  Never had a job
  Can't choose
  (ISSP Social inequality-IV, 2009)
- Concerning your family's total income over the past one month, which of the following is true?
  It was less than adequate for your family's needs
  It was just adequate for your family's needs
  It was more than adequate for your family's needs
  (Tanzania Kagera Health and Development Survey (LSMS), 2004)

are some nationally representative MyWorld surveys, for most part the large sample of MyWorld – over 7 million – has been possible because it has sought to be a truly worldwide survey and cover a wide range of countries, rather than to be nationally representative.

a) Perceptions data for guiding policy trade-offs

Perceptions data can play multiple roles in policy processes. Seaford (2013) defines six roles for subjective wellbeing data: to stimulate public debate, to draw attention to important issues which may otherwise be ignored, to inform the development of economic models, to influence the choice of other objective indicators, to input into a new form of cost–benefit analysis,[6] and to hold politicians accountable. These roles apply more broadly

to other types of perceptions data. We look in detail at the first two in this section, and touch on some of the other roles in Section 2b.

In the context of the post-2015 goals, there has been a systematic attempt include assessments and perceptions of different groups in the process of defining the new goals. But beyond this participatory decision-making process, there is greater scope for perceptions indicators to guide policy. While wellbeing data in developed countries have been used increasingly to support policies, they have been less used to drive the policy agenda (Seaford, 2013).

Perceptions data can help to drive the policy agenda by providing information on people's priorities. This can help analysing the interactions between different pathways from policies to wellbeing. For example, this data can help to make visible how policies in health, for example, are complementary with those in education or in security and stimulate public debate on the right mix of policies appropriate in a given country context. This also points to the trade-offs between conflicting objectives or a multiplicity of objectives that need to be prioritised.

Examples in the post-2015 context and beyond show how subjective indicators can fulfil this role. The OECD 'How's life' initiative and the Better Life Index[7] are examples of the latter. They identify 11 dimensions of material wellbeing and quality of life[8] for OECD countries and other major economies. The index is an interactive web-based tool that invites people to express the importance they give to different dimensions of wellbeing.[9] The key is that this exercise allows respondents to contrast countries according to the relative value they place on each dimension, rather than by having a static predetermined combination. Moreover, it allows for an illustrative comparative analysis of how these priorities differ across people from different countries, age and gender.[10] People's responses make it possible to understand how the high importance attributed to health, for example, is related to other dimensions of wellbeing and to the objective quality of health care systems in particular countries. In this sense, the work on new measures of progress, such as the Better Life Index, is useful in considering public priorities and strategies in policy-making (UN Task Team, 2013), identifying policies that would contribute more to overall wellbeing.

Similar exercises can be carried out with more traditional measures of wellbeing, to indicate what people value and to what extent, and to assess how well governments are responding to those trade-offs and fulfilling citizens' expectations of wellbeing. For example, Kroll (2013) proposes a measure of country performance in human development that combines the objective indicators of the Human Development Index (HDI) but weighted according to their contribution to SWB. This measure assesses the country correlation of the three HDI components (material conditions, health and education) and people's responses in a life-satisfaction question. By measuring which of the three HDI elements are more strongly related to life satisfaction, it becomes clear which countries are more successful in generating the goods that truly matter to people's wellbeing. Kroll finds significant heterogeneity in the importance of these three factors across 70 countries. For example, in Moldova, there is the highest correlation between SBW and income. In contrast, in Finland or Turkey, the contribution of income to SBW is low.

In health, use of the Quality Adjusted Life Years (QALY) and of Patient Reported Outcome Measures (PROMs) is a practical example of how these perceptions on trade-offs can be incorporated in the decision-making process of policy (Melamed et al., 2012). Pioneered in the UK, the exercise involves eliciting the 'patient's perspective' through thousands of surveys (known as the EQ5D[11]) asking people about their health. For each of the five health dimensions, individuals rate their health on the day when the questionnaire is completed – no problems, some problems or severe problems – and value each of the possible health outcomes.[12] For any given treatment, policy-makers need to know how much value is gained, for how long, and at what cost. One QALY is equal to one year lived in full health. A year in anything less than full health is valued at less than one, depending on the social value (derived from the PROMS process described above) attached to that particular outcome. Through this exercise, it is possible to assess the trade-off between health treatments that improve different aspects of health. For any given treatment, the number of QALYs gained (either through improvements in

6 The UK National Health Service (NHS) use of Patient Reported Outcome Measures (PROMS) as one input into measuring the performance of different health providers/services would be part of this cost–benefit analysis.

7 See http://www.oecd.org/statistics/howslife.htm and http://www.oecdbetterlifeindex.org/

8 Selected to match those proposed by Stiglitz et al. (2009), agreed by OECD countries and reflecting dimensions of wellbeing that are universal and relevant for all human beings rather than specific for a particular country. They are: community, education, environment, civic engagement, health, housing, income, jobs, life satisfaction, safety and work–life balance.

9 Users have to rate each of the 11 topics from 0 ('not important') to 5 ('very important').

10 Although to be precise, representative samples of the population would be required.

11 Each of its five sections corresponds to a different health outcome: mobility (ability to walk about normally); self-care (ability to look after oneself); usual activities (ability to perform usual activities); pain or discomfort; and anxiety or depression.

12 There are 243 different possible outcomes to the questionnaire (three possible outcomes on each of five dimensions of health).

quality or length of life, or both) can be compared with the total costs of the treatment (Melamed et al., 2012).

Trade-offs in budget allocations are especially visible. One of the successes of the MDG was to increase development aid from developed to developing countries (Clemens et al., 2007) but there are differences in what different groups of people perceive to be the best use for those resources. Using perceptions data to illustrate how people perceive trade-offs can be particularly useful in informing and monitoring budget decisions of donor countries, a comparison that is interesting in the post-2015 context where a similar exercise of assigning budget priorities may be needed. Leo (2013) investigates whether US aid is prioritising those issues that matter most to people in aid-receiving countries in Africa and Latin America. Regional Barometers ask what people believe are the most pressing problems facing their country. Only 16% of US assistance has been focused on what Africans cite as their top three most pressing problems (jobs, income and infrastructure) since 2002; nearly 60% has been targeted towards secondary or tertiary concerns (Leo, 2013). By contrast, Prizzon (2014) finds that, in Ethiopia, Cambodia and Zambia, citizens' preferences may not be necessarily aligned to the top priorities of their governments (infrastructure, energy and growth creation) but that the allocation of aid flows from donors does appear in those countries to be closely aligned to citizens' preferences.

In the context of defining the new set of development goals, information on people's preferences could be extremely helpful in establishing priorities for implementing the 17 goals so far proposed by the Open Working Group (OWG). Both Gallup and the WVS asked people to prioritise the MDG goals most important to them. In 26 countries of sub-Saharan Africa, according to the 2006 Gallup poll, reducing poverty and hunger were the highest ranked, followed by reducing the spread of HIV, and providing jobs, with a very consistent pattern by population subgroups but some variation by country, particularly in West Africa (Tortora, 2009).

Other interesting messages emerge from perceptions data that ask people to prioritise global and country goals. In the WVS, among other problems, 'people living in poverty' was overwhelmingly ranked top across countries as the most serious problem for the world. However, a lower share of people considered it the top country priority.[13] Moreover, when asked about whether the country should focus on solving its own problems or helping to reducing poverty in the world, most people think their country's problems come first.[14] Even in the most altruistic countries,[15] the majority of people would prefer to concentrate on their own country priorities rather than on global ones. The inclusion of sustainability, the environment and climate in the new set of development goals has broadened the scope of the SDG beyond the national spectrum. The information presented in the WVS aligns with the experience of MDG 8 'global partnership' (Kenny and Dykstra, 2013), suggesting that it is going to be hard to ensure global commitment to issues that are global in nature, or that require coordination and cooperation from various countries. Apart from the coordination issues and the alignment of political views by governments of different countries, ensuring public support within countries to tackle those issues will not be easy.

In sum, the value of subjective indicators is that they can make explicit the implicit assumptions that typically guide public debate, mainly around the issues that people consider important in improving their lives (Seaford, 2013; Rodriguez Takeuchi, 2014). This can be very helpful in informing policy debates but may not settle them (Seaford, 2013). Moreover, while it is important to know these overall preferences, perceptions data can also be used to assess trade-offs for different groups of people, for example to see the differences between what men and women, or younger and older people, think could most improve their wellbeing. In this sense, perceptions data can draw attention to important issues or groups of people which otherwise may be ignored, because they were not previously identified as a concern (Seaford, 2013).

b) Perceptions data for accountability

'To ask the people, with regularity, for their own thoughts strikes us as being both useful and a check on the claims of those in power' (Dionne and Mann, 2003)

There is no single agreed definition of accountability. It involves both answerability – the responsibility of duty-bearers to provide information about and justification of their actions – and enforceability – the possibility of penalties or consequences for failing to answer accountability claims, although this second aspect is often left aside in the understanding of accountability (McGee and Gaventa, 2011). There is still much debate about how to translate information, transparency and participation into accountability and improved outcomes, and whether, in fact, there is a clear positive connection (McGee and Gaventa, 2011). Nevertheless, the very nature of perceptions data – asking people what they think on a range of dimensions of their wellbeing – has the potential

13 63% consider it a global priority and 58% a national one.

14 In a scale from 1 to 10, the average score was 7.8. About 39% of respondents marked 10 ('it is a top priority to solve my own country's problems').

15 Finland, Iran, Italy, Malaysia, Norway, Mali, Mexico, Spain, Sweden and Switzerland, according to the survey results.

to generate greater accountability. This works in two ways: first, by opening the debate around measuring and acting on what is truly important to people; and, second, by increasing the availability of information and hence, opportunities to track progress.

Accountability starts with the creation of systems that allow it. Making sure to measure things that people actually think are important and relevant is a first step. There is a wide debate on whether social accountability can occur by making rules, procedures and plans transparent and open for discussion before they are implemented (McGee and Gaventa, 2011). In this sense, perceptions data could be used to open this discussion and even influence the choice of indicators for tracking outcomes that governments will be held accountable for until 2030. Even if perceptions data are not directly incorporated in the monitoring indicators of the SDG, they can influence the choice and design of other indicators and highlight important issues expressed by people which may otherwise be ignored.

This exercise may also reveal that commonly used objective indicators may not be reflecting the appropriate issues. For instance, there often seems to be a marked distance between standard measures of important socioeconomic variables like economic growth, inflation and unemployment, and widespread perceptions of countries' development progress (Stiglitz et al., 2009). In some countries, these gaps have undermined confidence in official statistics,[16] and in others they have resulted in a partial if not misguided analysis of countries' and people's situations.

'All institutes worldwide knew GDP was rising in Tunisia and Egypt. They knew what 11 million Tunisians and 80 million Egyptians were buying and selling — but they didn't know what they were thinking. As a result, revolutions in those countries came as a shock. The UN didn't see those revolutions coming, neither did the WEF nor the World Bank. The U.S. spends tens of billions on intelligence — and it missed those revolutions too.' (Gallup, 2012)

Perceptions data can thus contribute to the social relevance of indicators by helping to discern the most relevant issues for policy-makers to address – for example, by identifying objective indicators that better correlate with people's subjective wellbeing. Seaford (2013) points out that to measure a country's economic prosperity, for example, GDP per capita (mean GDP) could be replaced by median income, which takes into account distributional issues and has a greater impact on wellbeing than average (mean) income. Similarly, looking at the relationship between, say, perceptions of safety and a range of indicators on safety and security may be used to inform the selection of an objective indicator to monitor a post-2015 target on security. Moreover, perceptions indicators are socially relevant because they collect valuable information on domains of wellbeing that are intrinsically subjective (Gough and MacGregor, 2009) and thus difficult to measure with objective data.

In this way, perceptions data can help to make governments accountable for social progress, defined in terms of contributors to multidimensional wellbeing, rather than a narrow set of outcomes. Using subjective data then might improve the accountability of global and national decision-makers with respect to the overall aim of the post-2015 goals: to improve people's lives.

In the specific case of indicators of governance, perceptions data can be used to hold governments accountable for democratic progress. While many objective measures of governance tend to focus on formal procedures or institutions (e.g. whether and how often elections are held, the formal separation of powers, the existence of an anticorruption commission), governance progress requires measures along various dimensions (Foresti et al., 2014) that underpin the relationships between the state and society. Perceptions data may add the perspective of those people whom governance is supposed to benefit.

Corruption is a key example. Asking people whether they think officials who commit crimes go unpunished, or finding out the number of firms that report that informal gifts or payments are expected to obtain services (UNDP, 2014), provides information on how people experience corruption rather than just about prevalence levels. Other perception measures important in terms of accountability and governance include the extent to which citizens feel they are able to participate in the selection of their governments, have freedom of expression and live in a country where media are free and offer a variety of views (Thomas, 2009). If these measures are properly integrated, they give a critical voice to people who previously have had limited or no voice and had therefore been excluded from accountability mechanisms. The important point here is how, and perhaps if, the data are used. Data alone are not an accountability mechanism. But if the data are appropriately used by, for example, free media to raise issues highlighted in people's perceptions then this can make a difference to how politicians act and respond to people's needs.

16 For example, in France and in the United Kingdom. only one third of citizens trust official figures, and these countries are not exceptions (Stiglitz et al., 2009).

Perceptions data, especially in combination with new methods of data collection, can also fill another important gap in measuring outcomes to track progress towards the post-2015 targets: the availability and timeliness of data. Increasing the frequency and coverage of data is a key component of the data revolution. This is crucial not only to ensure accuracy of the statistics used to show progress (or lack of it), but also to understand the impacts of shocks, sudden changes in wellbeing and instability situations. In particular, perceptions data can act as social monitors and point to changes in trends, before those are reflected in other types of statistics. For instance, the subjective food insecurity scale of the Food and Agriculture Organization (FAO) shows the potential of such indicators to detect early onset of malnutrition crises. Unlike a measure of food availability from agricultural production or trade, or anthropometric indicators such as underweight or stunting prevalence, which evolve only after malnutrition becomes manifest, a perception-based food insecurity index can be used to identify vulnerable populations in a timely way (Ballard et al., 2013) and even help to identify seasonality patterns (Headey and Ecker, 2013).

This same idea was behind the Social Weather Station (SWS) project, one of the first perceptions surveys in the developing world – the idea that 'surveys can serve like observation posts to monitor social conditions, much as meteorological stations monitor weather conditions' (Mangahas, 2013). As such, perceptions data offer a unique opportunity to capture the evolution in people's perceptions and potentially serve as an 'early warning system' (Alkire and Samman, 2014)[17] for the need for policy intervention. In fact, an exploration of the SWS self-reported poverty data and the official poverty statistics of the Philippines reveal that, while self-reported poverty and official poverty statistics differ greatly in terms of levels, the trends they describe are broadly consistent (Alkire and Samman, 2014).

Although some headline objective indicators such as GDP are tracked very regularly in more developed countries, in some other countries equivalent data can be unreliable (see for example Jerven's (2103) argument about core statistics in Africa) or much delayed. With respect to MDG indicators, for example, the most recent global report includes data up to four years old: 'The poverty figures go back to 2010. We're now in the middle of 2014, so that's quite frustrating' (Keiko Osaki-Tomita[18] in SciDev, 2014).

Increasing the frequency of surveys can be costly and highly demanding for NSOs with limited capacity. More complete sources of data such as censuses tend to be very infrequent, often conducted once every decade. Household surveys have been rising in number but are still relatively limited, especially in low-income countries. Chandy (2013) for example finds that, in low-income countries, household surveys used to measure monetary poverty have been carried out on average just four times since 1980. Similarly, Demographic and Health Surveys (DHS), key surveys used for tracking MDG, have been carried out in 89 countries from the late 1980s to date. However, of those 89 countries, 26 have only been surveyed once.[19] In contrast, current perceptions data surveys are collected more frequently and thus have the potential to be collected in a timely manner and frequently in a wide range of countries. The SWS project has tracked Filipinos' perceptions at quarterly intervals for over 30 years. The GWP, for example, has a coverage of 160 countries and is fielded annually in 93 of them (Alkire and Samman, 2014). This may be partly because collecting perceptions data is simpler and less costly (see Ballard et al., 2013; Headey and Ecker, 2013, in the context of food security data). However, addressing some of the reliability issues (as discussed in Section 4) requires the application of methodological rigour, interviewer training and ensuring the appropriate survey protocols, all of which can increase the length and cost of perception surveys in the future. The cost and time involved in collecting perceptions data is also contingent in the specificity of the instruments and ways of data collection.[20]

Finally, a combination of big data and subjective data can be used to extrapolate hard data, for 'nowcasting'. This is immediate-term ('now') forecasting on the basis of real-time data flow (World Bank and SecondMuse, 2014). This technique is used by statistical offices to fill in the gaps of series with low frequency, such as in forecasts of food prices or disease onset using Google trends analysis. For example, in the context of perceptions data, a joint World Bank and UN Global Pulse pilot explored the potential of social network content (from Twitter) to analyse public perceptions of a gas subsidy reform in El Salvador. The study used text analytics to see if the results from the social media analysis, which can be tracked with high frequency, matched public opinion as measured through household surveys conducted before and after the reform. Preliminary results confirmed that Twitter data provided a useful complement to analyse the public perception of a policy reform (World Bank and SecondMuse, 2014) and thus can be used to capture the evolution in people's perceptions, potentially serving as a warning system, as mentioned above. This could point to the need for policy intervention before the full household survey comes into place. A similar case could be envisioned for example for a UNICEF

17 Cummings (2009) suggested a similar argument for subjective wellbeing data.

18 Chief of the demographic and social statistics branch of the UN's Statistics Division and lead author of the 2014 UN MDG report.

19 http://dhsprogram.com/data/available-datasets.cfm

20 The use of new technologies may help to reduce costs and time, although this discussion falls beyond the scope of this paper.

project monitoring social media and public blog posts to track parents' attitudes towards vaccination, particularly vaccine hesitancy in Eastern Europe. High-frequency data on these attitudes would be very valuable in planning intervention before the programme comes to an end, when very little could be done to change adverse perceptions of child vaccination.

Filling measurement gaps

This section concentrates on the potential of perceptions data to guide policy and monitor progress for particular post-2015 targets where existing data gaps are sizeable. Based on an analysis of the existing proposals for post-2015,[21] where use of perceptions data has been suggested as part of the measurement strategy, and the assessment of the UN task team on indicators (UN Task Team, 2013: ix), we identified three areas where this type of indicator can be of particular relevance to supplement objective indicators and provide a fuller picture. These are: a) gender and social norms, b) governance and service delivery, and c) violence and security. This section this focuses on how perceptions data can expand the possibilities for measurement.

a) Gender and social norms

Goal 3 of the MDG is to 'Promote gender equality and empower women'. The main target focuses on eliminating gender disparities in education, with complementary targets on gaps in wage employment in the non-agricultural sector and on political representation. Traditional gender indicators measure progress on outcomes of particular relevance for women or outcome gaps between men and women. Examples are maternal mortality, early marriage and female genital mutilation (FGM) rates. The second type of indicator involves disaggregating existing indicators by gender. MDG gender indicators, for example, are of this second type: the share of women in national parliaments, gender parity in education, and the share of women in non-agricultural wage employment. A third type of indicator relates to the commitment of national governments to ensuring gender equality. The Inter-Agency and Expert Group on Gender Statistics (IAEG-GS) measures 11 of those commitments through the existence of the relevant legislation or ratification of international conventions and commitments.

Most traditional gender indicators, whether outcome-based or procedural, are based on objective data. The set of minimum gender indicators recommended by the IAEG-GS includes 52 indicators;[22] none is based on perceptions data. Nevertheless, in the context of the new set of development goals, achieving gender equality and empowerment may require not only tracking gender gaps in outcomes in 2030, but also monitoring the evolution of social norms. Social norms are defined as 'informal and formal laws, beliefs and practices that help to determine collective understanding of what are acceptable attitudes and behaviours' (Harper et al., 2014: 2), and are a key component of progress towards gender equality and empowerment.

In the post-2015 discussions, gender has featured strongly in goals proposed by both the HLP and the OWG. In the OWG proposal, the goal is stated as to 'End all forms of discrimination against all women and girls everywhere', and the HLP report states that 'these barriers (to opportunities for women and girls) can only be removed when there is zero *tolerance* of violence against and exploitation of women and girls' (HLP, 2013: 17, italics added). Some of these forms of discrimination will be reflected in objective indicators, and ratios between men and women are useful to capture them. But some other forms of discrimination, and the zero tolerance described by the HLP, are harder to capture.

Perceptions data can help to measure the evolution of social norms that drive unequal outcomes for women through their life-cycle in domains such as education, health, political representation and labour markets (Branisa et al., 2009 and OECD, 2010 (*Atlas of Gender and Development*) cited in Harper et al., 2012). For example, studies by Hussein (2010) in the UK, or Yirga et al. (2012) and de Cao and Lutz (2014) in Ethiopia, try to understand the beliefs behind the practice of FGM rather than its prevalence. They found that FGM can often be unquestioned by women themselves, and perceived as a traditional requirement, despite the pain involved, to avoid social isolation. This type of information may be helpful to understand how, beyond legal frameworks, informal social norms underpin unequal or discriminatory outcomes for women.

Indicators to measure social norms using perceptions data have started to appear. The OECD's SIGI (Social Institutions and Gender Index)[23] launched in 2009 is the first cross-country measure of social norms, looking at the *de jure* (legal) as well as the *de facto* (actual) situation of discriminatory social institutions in five categories: family code, civil liberties, physical integrity, son preference, and ownership rights. In SIGI, both attitudinal and prevalence data are essential to capture the *de facto* situation of

21 See http://tracker.post2015.org/

22 See http://genderstats.org/

23 http://genderindex.org/

women. Based on SIGI, Harper et al. (2014) propose post-2015 targets measured by changes in social norms as signals for growing empowerment of women and girls. The Women's Empowerment in Agriculture Index (WEIA),[24] developed by the International Food Policy Research Institute (IPFRI), is another pioneering gender-related index that includes perceptions data at the core of the measurement of several of the indicators comprising the 'Five domains of empowerment (5DE)' index.[25] As in SIGI, a large number of indicators that comprise the index are based on perceptions data (see Alkire et al., 2013).

Indicators based on social perceptions (held by both men and women) are at the base of various proposals to measure a gender goal post-2015. Table 1 presents selected examples showing how both types of indicators could be used in various sub-categories of women's empowerment. For instance, measuring a target on sexual and reproductive health can be done through prevalence rates of unmet need for contraception, but also through responses to questions about perceptions towards sexual attitudes (in particular about the right to refuse sex with a partner). These areas are not necessarily covered in the MDG framework or all featured in the SDG. However, most of these examples have been proposed for monitoring the post-2015 goals. A few additional indicators currently used for measurement were added to the table for indicative purpose. There is not necessarily a one-to-one correspondence between the perception and the objective indicators, but the contrast is useful to note with the different information that perceptions or objective indicators provide.

b) Governance

Governance was a key area left out of the MDG. In contrast, the importance of good governance in itself and towards the fulfilment of goals for other sectors has been recognised in proposals for post-2015 targets. The HLP's report (HLP, 2013) has one of its 13 goals as 'Ensure good governance and effective institutions' and proposes the following five targets:

1) Provide free and universal legal identity, such as birth registrations.

2) Ensure that people enjoy freedom of speech, association, peaceful protest and access to independent media and information.

3) Increase public participation in political processes and civic engagement at all levels.

4) Guarantee the public's right to information and access to government data.

5) Reduce bribery and corruption and ensure officials can be held accountable.

The Open Working Group's latest proposal (2014) has a narrower focus, with goal 16 including some aspects such as building 'effective accountable and inclusive institutions at all levels'. Although each of these goals is accompanied by more various targets, the detail of the specific indicators and measures that would be used to track these targets has not been announced. Measuring these issues appropriately will be key to understanding the extent to which countries are progressing (or regressing) relative to their targets.

To date, measures of governance often focus on the form of government rather than its performance, for example, on the existence of particular legislation or policy and the ratification of conventions (Foresti et al., 2014). Governance is also closely connected with service delivery. While the quantity aspect of service delivery is well suited to objective measures, issues to do with the quality and access to these services are often better suited to subjective measures. The post-2015 discussion has focused on refining and expanding these measures of service delivery but still along objective lines. An optimum situation would be the use of both objective and subjective measures to understand both the form and function of governance, and whether it is delivering what people want. Table 2 shows some of the different possible perception measures that could be used, compared to objective measures in similar areas.

There are however, a few considerations when using perceptions data on service delivery. Perception measures around education, for example, may be unreliable if we are dealing with first-generation learners. Parents who are themselves non-literate or who have not been to school will find it difficult to pass qualitative judgement on the kind of schooling being provided to their children. The media can play an important role in shaping perceptions. If mass media are relatively free from state manipulation, they can act as an accountability system both to inform the public of important issues and for members of the public to raise these issues themselves. Where the state has control of the media, they can be used to manipulate people's perceptions of the quality of governance. It should be noted that objective measures are as much at risk of government manipulation, for example by adopting certain 'forms' that lead to 'better' scores on an indicator, but without real change in the underlying aspects of governance (also known as isomorphic mimicry) (Foresti at al., 2014).

Framing effects may be important in evaluating perceptions of service delivery. The phrasing of single questions can have a significant impact on people's responses. This is demonstrated in perceptions data around the UK's National Health Service (NHS), or in use of words such as 'assistance to the poor' versus 'welfare' (Bartels (2003); as discussed in Section 4b below). In the UK, when asked in 2011 about their local NHS provider,

24 http://www.ifpri.org/publication/womens-empowerment-agriculture-index

25 The five domains are: production, resources, income, leadership and time.

Table 1: Selected gender and social norms indicators

Target: 'End all forms of discrimination against all women and girls everywhere'		
Area	**Perception indicators**	**Objective indicators/self-reported objective proxy**
Sexual and reproductive integrity/autonomy	Percentage of women and men who think a woman can refuse sex with her husband/partner under any circumstance.[U]	Percentage of married women with an unmet need[***] for family planning.[H]
Freedom from violence	Percentage of women and men who think that a husband/partner is never justified in beating his wife/partner.[H,U]	Lifetime prevalence of domestic violence.[H,S,U] Percentage of women who have undergone female genital mutilation.S,U
Decision-making ability about land, assets and resources	Extent to which the individual feels he or she can make his or her own personal decisions regarding the aspects of household life[*] if he or she wanted to.[W] Relative Autonomy Indicator.[**] [W] Percentage of people who think important decisions in the household should be made by both men and women, by sex.[U]	Self-reported questions on decision-making and women's participation in decision-making and control over earnings in households. For example: who decides how to spend money (the wife? the husband? both?)[H] Self-reported questions on decision-making and woman's participation in decision making regarding economic activities. For example: If an individual participated in the activity,[****] how much input did the individual have in making decisions?[W] Proportion of adult population owning land, by sex.[U] Proportion of population with access to institutional credit (other than microfinance), by sex.[U]
Participation in political and civic life	Proportion of those of voting age who agree or strongly agree that, on the whole, men make better political leaders than women.[H]	Score based on legal restrictions or discriminatory practices affecting women's access to public space.[*****] Percentage of women in national parliament.[U]
Equal value to boys and girls	Percentage of respondents who agree that 'education is more relevant for boys'.[H]	Missing Women indicator (gender bias in mortality due to sex-selective abortions, female infanticide or insufficient care given to baby girls).[H,S]
Time	Rank their level of satisfaction from 1 = not satisfied to 10 = very satisfied, with the time available for leisure activities such as visiting neighbours, watching TV, listening to the radio, seeing movies, or participating in sports.[W]	Female-to-male ratio of average time devoted to household activities/unpaid domestic work.[H,U] The female-to-male ratio of total workload (both paid and unpaid work).[H]
Support services and justice for women and girls	Proportion of the population who feel safe walking alone at night in the area where they live, by sex.[U]	Proportion of women over 15 years old subjected to physical or sexual violence in the past 12 months who reported it to the justice system.[U] Proportion of law enforcement professionals who are women (including judges and the police).[U]

Sources: H - Harper et al (2014), W - WEIA, S - SIGI, U - UNWomen (2013)

[*] Activity areas: (a) which inputs to buy, (b) which types of crops to grow for agricultural production, (c) when to take or who should take crops to market, and (d) whether to engage in livestock-raising.

[**] Formed of three questions: 1. My actions in [activity area] are partly because I will get in trouble with someone if I act differently; 2. Regarding [activity area] I do what I do so others don't think poorly of me; and 3. Regarding [activity area] I do what I do because I personally think it is the right thing to do.

[***] Women who are fecund and sexually active but are not using any method of contraception, and report not wanting any more children. This definition points to the gap between women's reproductive intentions and their contraceptive behaviour and thus it incorporates a perceptions component ('not wanting more children'). We place this under the self-reported proxy category.

[****] About food crop farming, cash crop farming, livestock raising, and fish culture.

[*****] For example, the restrictions on women's choice of domicile, restricted ability to visit family and friends, requirements for husband's approval to apply for a passport or widespread threats of political violence.

Table 2: Selected governance and effective institutions indicators

Target: 'Ensure good governance and effective institutions'		
Area	**Perception indicators**	**Objective indicators/self-reported objective proxy**
Government effectiveness	Satisfaction with education system	Quality of public schools Increase in tax revenue as a proportion of GDP
Control of corruption	How many government officials do you think are involved in corruption? Proportion of public who believe they can receive timely services without paying a bribe	Government efforts to tackle corruption Reduction in number of people who report paying a bribe
Voice and accountability	Satisfaction with democracy	Freedom of the press Open Budget Index score

Sources: Kaufmann et al. (2009) based on various sources; Foresti et al. (2014) based on various sources

77% of respondents agreed that they were being provided with a good service. When asked about whether the NHS delivers a good service nationally, only 63% of the same respondents agreed that the service was good. When asked about the government's NHS policies, approval levels dropped to 28%. This trend has been maintained since data were first collected in 2003 when the figures were 76%, 48% and 30%, respectively (Ipsos Mori, 2012).

c) Violence and security

The MDG did not include any goals or targets around violence and insecurity, although momentum is building for their inclusion in the post-2015 agenda. Freedom from fear, conflict and violence was one of the five transformative shifts to drive the universal development agenda outlined by the HLP (2013). It is also a basic human right and a core element of wellbeing. The OWG's proposal (UN, 2014a) includes targets on eliminating all forms of violence against girls and women, significantly reducing all forms of violence and related death rates, and ending all forms of violence and torture against children.

Clearly, some of these issues can be measured with objective indicators. Traditional indicators include homicide and crime rates and clashes between armed groups (legal or illegal). These are broadly comparable across countries and offer insights into the level of security risk posed to the general population. However, people's perceptions of violence and insecurity do not always correlate with objective measures. Stiglitz et al. (2009) report that there can be significant differences between actual and perceived rates of violence and insecurity, with the media potentially playing a significant role in distorting people's views. When analysing the victimisation rate compared with the fear of crime in OECD countries in 2005, no correlation was found between the two. Similarly, Garrett and Ahmed (2004) found that in the city of Dinajpur in Bangladesh, despite the high crime rate, only 10% of men and women reported that they did not feel safe.

In situations in which social norms dictate that certain types of violence are acceptable, people may underreport it. A strong example of this is around FGM. In countries where this practice is accepted and even desired, this type of violence will be underreported by girls (Hussein, 2010; Yirga et al., 2012). Similarly, some people may believe that certain levels of domestic violence are to be expected and therefore do not perceive them as harmful. This suggests that we should not rely solely on perceptions data but should use them in conjunction with objective data.

Large-scale objective measures will give only a surface-level understanding of the volume of violence at best, whereas subjective measures can help elucidate the underlying causes of this violence by asking people more specific questions around how this violence is experienced. Subjective indicators are more nuanced and offer a deeper understanding of how violence and insecurity are perceived by people and can help to understand underlying dynamics in fragile and conflict-affected situations (Mallett, 2012). In the Bangladesh case above, the role of *mastaan* (local strongmen) who offer people 'protection', was a possible cause for the mismatch between objective and subjective measures. Nevertheless, this came at the cost of those most marginalised, or not in the favour of these strongmen figures, pointing to some inequalities in the provision of security. Significant mismatch between perceptions and objective measures can highlight significant problems in a society, for example lack of access to justice or to the institutions in charge of protecting civilians and delivering security.

A key component of the concept of security is how safe people perceive their environment to be, and how their wellbeing interacts with this perception. For example, Møller (2005) finds that the perceived likelihood of victimisation and concern about personal safety has a stronger negative impact on life satisfaction than actual levels of victimisation.

Table 3: Selected indicators to measure violence and insecurity

Target 1: 'All social groups are free from violence and insecurity'		
Area	**Perception indicators**	**Objective indicators/self-reported objective proxy**
Equal justice	Do the police treat people equally? (piloted by the Vera Institute of Justice)	Percentage of police, prosecutors and judges that are women
Safety	Do you feel safe walking alone at night in the city or area where you live? How safe do you feel now moving to other places e.g. markets or town?	Political stability and absence of violence score
Trust in police	How much do you trust the police?	
Police capabilities	How easy or difficult is it to get help from the police?	Number of convictions over number of police

Sources: Saferworld (2013) based on various sources; SLRC surveys

There are a variety of proposals for subjective indicators that could be introduced in the post-2015 framework. Saferworld (2013), which has one of the most detailed proposals concerning violence and security, gives several examples of existing subjective indicators that could be incorporated to track two proposed targets: a general target around violence and insecurity (Table 3), and specifically gender-based violence (Table 4). The Secure Livelihoods Research Consortium (SLRC)[26] surveys in conflict-affected countries also presents some valuable examples. Again, there is not a one-to-one correspondence between the objective and perceptions indicators in the tables, which is indicative of the different aspects captured by each type of data.

In sum, the review of these three areas shows that perceptions data could serve as a useful complement to objective indicators of progress. Objective indicators are useful but they have their limits, and sometimes additional perceptions indicators are required (Veenhoven, 2004; Gough and McGregor, 2009). For example, in the case of gender, the reduction in gaps between men and women manifests only partly improvements to gender empowerment and equal opportunities. Data reflecting social beliefs on gender roles and accepted behaviour can also point to important advancements in achieving gender equality. In this sense, subjective perceptions are very useful to assess policy success (Veenhoven, 2004) but less effective for assessing why a specific policy worked and how to generate such changes (Hulme et al., 2014).

Table 4: Selected indicators to measure progress on eliminating violence against women and girls

Target 2: 'Violence against women and girls is eliminated'		
Area	**Perception indicators**	**Objective indicators/self-reported objective proxy**
Violence against women	Percentage of population believing that a husband is justified in hitting or beating his wife/partner Percentage of women vs. men who believe that the police would respond if they reported a crime (piloted by Vera Institute of Justice)	Number of recorded rapes per 100,000 women and girls Homicides of females per 100,000 females

Source: Saferworld (2013) based on various sources

26 http://www.securelivelihoods.org/

Why not perceptions data? Common but differentiated challenges of perception surveys

Interest is increasing from governments in understanding, collecting and using perceptions data. A wider range of research helps to understand the main concerns and issues to consider here. The data also have some limitations, which we discuss in this section. Thus, it may be essential to complement perceptions data with objective measures or administrative records from other sources.

The issues discussed in this section are not specific to the post-2015 context, but rather relevant in general to the collection of perceptions (and other survey) data. Because perceptions data are collected through individual or household surveys, key issues in using this type of data relate to survey methods and their inherent limitations. Error may arise from well-known flaws affecting surveys. This does not imply mistakes, in the colloquial sense of the word, but rather deviations from what is desired in the survey process (Groves et al., 2009). In broad terms, survey errors can be divided into two groups: errors of observation or measurement; and errors of non-observation, which relate to the process of obtaining statistics from a sample of the population (Figure 1). We touch on some of the issues of representation first, then focus on a selection of measurement error sources that could be important in the context of collecting perceptions data, particularly because of the nature of asking people to self-report their opinions. This review is not intended to be exhaustive but should indicate the main issues of reliability of perceptions data.[27]

a) Representation: whose responses are represented?

First, as in any survey, possible concerns include: the representativeness of the sample, sample numbers, low response rates and accessibility issues when infrastructure is limited.[28] Adopting a rigorous sampling mechanism is a necessary step for any survey, as is reporting on sampling errors, to have a sense of how close the sample is to the population of reference.[29] Nevertheless, in comparison with some large-scale multi-topic household surveys currently used in the monitoring of the MDG, existing perception surveys tend to have much smaller sample sizes. For example, the sample size in the Gallup World Poll is around 1,000 people and up to 2,000 in large countries such as China and Russia.[30] The surveys also have sample sizes of about 1,000 people in each country, although in some exceptional cases larger samples have been drawn.[31] In contrast, a standard DHS has a sample of between 5,000 and 30,000 households.[32] Multiple Indicator Cluster Surveys (MICS), as used by Unicef, also tend to have large samples; the most recent MICS survey for Vietnam for example, interviewed close to 10,000 households and 10,000 women.[33] If perceptions surveys became part of the

27 For a detailed review, see Groves et al. (2009) or Krosnik and Presser (2009), among others.

28 See, for example, Traugott (2003) for some of the issues concerning electoral polls in the US, and Herbert (2013) on perception surveys in conflict-affected countries.

29 See http://www.aapor.org/Best_Practices1.htm#.U9Z1_PldU9Y for best practices in perception surveys.

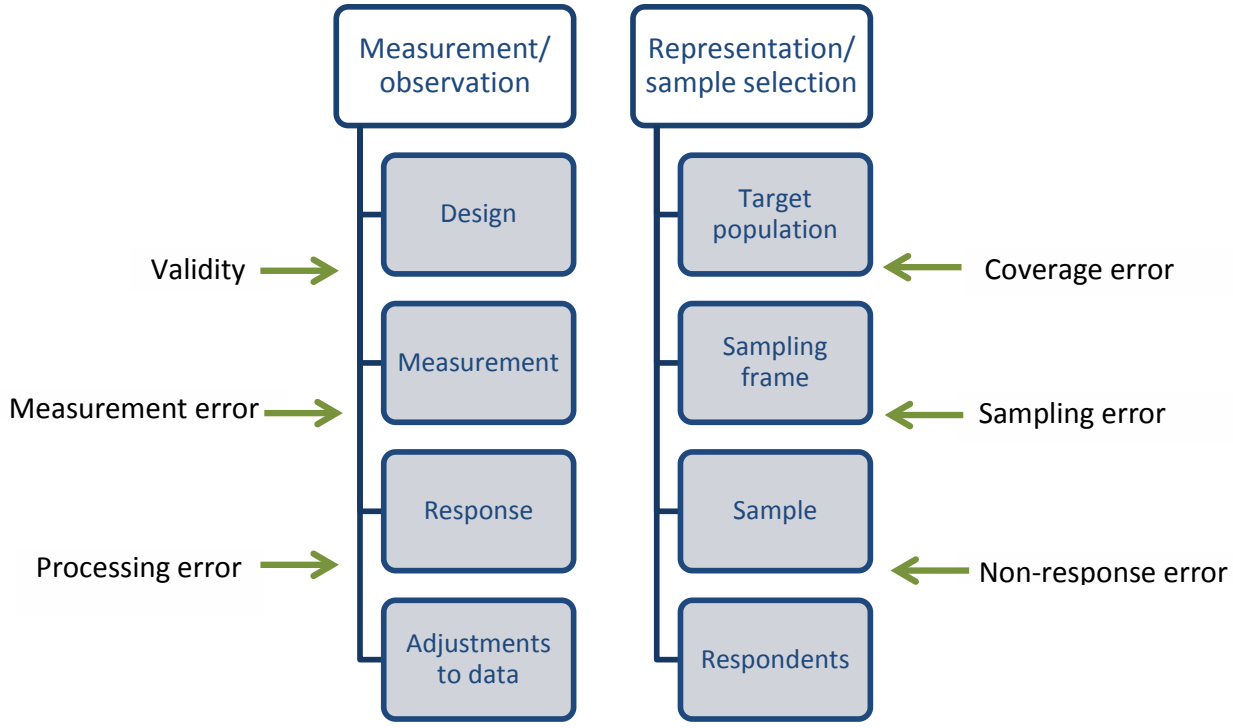
30 http://www.gallup.com/poll/105226/world-poll-methodology.aspx

31 http://www.hks.harvard.edu/fs/pnorris/Acrobat/Global%20Communications/Technical%20Appendix%20%20B%20List%20of%20Countries.pdf

32 http://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm

33 http://www.childinfo.org/files/Viet_Nam_2013-14_MICS_KFR.pdf

Figure 1: Errors in the process of obtaining estimates from surveys



Source: Adapted from Groves et al. (2009)

SDG monitoring, or perceptions modules were consistently added to existing larger household surveys, perceptions data might become much more representative.

Even with the bigger sample sizes of traditional household surveys, there may be issues when trying to gather data for specific population groups likely to be unrepresented. Moreover, there is a bigger point related to a key focus of the SDG in finishing the MDG task by reaching the hard-to-reach and 'leaving no one behind'. Are surveys really going to be the best way to gather the experiences of these groups? This has been highlighted as problematic for instance in the case of those living in informal settlements of urban areas, older people and people with disabilities (Carr-Hill, 2013; Lucci and Bhatkal, 2014; Samman and Rodriguez Takeuchi, 2014). Moreover, the possibility of producing representative results for population subgroups beyond gender, place of residence (rural/urban), and sometimes age and national sub-region, gets smaller with lower sample sizes. Surveys need to be specific when reporting results about the representativeness of their data, so that it is clear what level of disaggregation is possible. It may be necessary to increase sample sizes to be able to provide sufficient information that does not mask the heterogeneity of points of view.

In some cases, acting on aggregate preferences may not be desired for equity reasons. The analysis of subjective data may contribute to post-2015 equity objectives by highlighting how population subgroups may benefit from the different goals and may be unevenly affected by different policies. However, the variable effect of different priorities and trade-offs on different subgroups of population, according to age, gender, socioeconomic position, culture, and many other characteristics (Bellani et al., 2013),[34] may also pose an important challenge in using these data. In particular, there could be systematic biases underlying perceptions data. For example, the well-known 'U' relationship between life satisfaction and age (Blanchflower and Oswald, 2011), the gender differences in reported life evaluations (e.g. Boarini et al., 2012), or the higher dissatisfaction with democratic institutions reported by young people (Bergh et al., 2014) may point to an important policy issue to be solved. Further, unless we know more about the reasons for the systematic differences, there is the risk of informing policy on the basis of incorrect information.

34 In principle, 'they can vary between individuals almost continuously, though for policy purposes the identification of discrete target groups may be easier to interpret and more useful for the development of policy' (Bellani et al., 2013).

A separate issue concerns gathering perceptions in areas where data collection is difficult because it touches on sensitive issues. Many of the questions proposed to date for the areas of gender norms, violence and security and governance may be difficult to gather for this reason. Response rates appear high for some perception-based questions. In the GWP and the WVS for example, (objective) income questions had between 10 and 100 times higher non-response rates than did subjective wellbeing questions, depending on the country (Smith, 2013, cited in OECD, 2013). For other types of perceptions, item non-response may be a bigger problem, although the evidence is limited. For example, non-response could be higher for questions or items that are difficult to comprehend, offensive to participants, or concern sensitive issues, and also in surveys that require longer interview times. Non-response in self-reported questions about living conditions has been analysed (for example, Riphahn and Serflinf (2002), for income), and direct survey questions on sensitive issues are likely to result in low participation and misreporting of answers, particularly for those with something considered 'embarrassing' to report (Tourangeau and Yan, 2007; Tourangeau et al., 2010). However, that evidence is limited to self-reported proxies of objective outcomes (for example drug use, voting, income), which fall beyond our definition of perceptions data. The extent of item non-response for other perception questions needs to be examined further.

Guidance and best procedures to minimise question and survey non-response have been developed for sections of household surveys touching on sensitive issues, and could be adapted to surveys that are exclusively dedicated to perceptions. For instance, since 1998-99, DHS set the task of developing a standardised approach to measure domestic violence, accompanied by guidelines on the ethical implementation of the questionnaire.[35] Some of the criteria are, for example, to ensure privacy and anonymity of the responses, having an interviewer of the same sex as the respondent, asking sensitive questions towards the end of the survey when rapport has been built up, or making available information on organisations that provide services or referrals to victims of domestic violence. These measures are also helpful to solve some of reliability issues discussed below.

b) Reliability: how accurate are people's responses?

The next issue is to ensure that the responses produce reliable information, indispensable for tracking the post-2015 targets. Polls and pollsters often face a lack of trust and are especially criticised when opinion moves in the 'wrong' direction (Dionne and Mann, 2003), which could be contrary to either individual beliefs or the predominant socially accepted view. Questions about the independence of the pollsters are often raised, but a more fundamental question concerns the reliability of self-reported information and the underlying questionnaire used to obtain such information. People may under- or over-state their responses or report false information, either intentionally or unintentionally, and such misreporting may bias subjective indicators. Unlike with objective indicators, cross-validation is not possible by simple observation of living conditions or administrative records. The way in which questions are asked, and the questionnaire itself, are perhaps the most important considerations in reducing measurement error (Krosnik and Presser, 2009).

We highlight four key problems that can lead to inaccurate information on perceptions:

- **Recall bias**: this occurs when individuals make mistakes in recalling past events. For example, when asked about how many times during the past week/month they experienced a certain feeling (happiness, anxiety, etc.). This may limit reliability of answers requiring participants to compare current and past perceptions.
- **Social desirability bias**: this occurs when participants respond with answers thought to be more socially acceptable or desirable, or when they restrict their answers for fear of repercussions. For example, in conflict situations, people may fear reporting their actual perceptions of the state (Herbert, 2013). Social desirability may also be present when discussing illegal topics (e.g. related to sexual behaviour or drug use). This type of bias may be particularly relevant in highly politicised societies or where freedom of expression is limited.
- **Framing effects**: this occurs when the way in which questions are asked affects the answers given. This type of bias can interact with social desirability bias. Take for instance the US General Social Survey (1984, 1985 and 1986): respondents were asked whether they thought their country was spending 'too much, too little, or about the right amount' on each of a variety of government programmes. While only a quarter of the respondents each year said that too little was being spent on *welfare*, close to 65% said that too little was being spent on *assistance to the poor*. These differences are attributed to the different connotations of the terms 'welfare' and 'assistance to the poor' in the US (Bartels, 2003). Questions that 'lead' the answer may suffer from both framing effects and social desirability issues. For instance, questions such as 'Do you think that...?' and 'Don't you think that...?' pull 'yes' for an answer (Barker et al., 2002). On the other hand, in some cases it may be useful to include such implicit premises of accepted behaviour to encourage the respondent

35 Given that they are part of the same survey, as is often the case with questions on *experience* of domestic violence and *perceptions* of domestic violence, this guidance is also applied to the perception-questions section of the survey.

to speak honestly. For example, studies on sexual behaviour often ask 'How old where you when…?' rather than 'Did you ever…?' (Barker et al., 2002). The order of questions or even whether the respondent is alone or in company may lead to different results. Sensitive questions tend to be left for last when there has been time to build rapport between and the interviewee feels more positive about expressing her true opinion.

- **Proxy respondent**: when a question is asked not directly to the person involved, but to a nominated household member, for example the head of household, the mother (of a child), or any available adult in the household at the time of interview. Bardasi et al. (2011) show that this can greatly affect estimates derived from household surveys.

Measurement error may be minimised through good evidence-based questionnaire design. For example, this can include: using recording sheets – forms where respondents write down the information over the course of a period – to reduce recall bias; having an interviewer of the same gender to build rapport; using anonymous responses, as for example in MyWorld, to encourage honest answers about sensitive issues; introducing probing questions and using plain language to avoid confusion and misreporting.

There are particular problems with self-reported proxies of objective indicators, likely to result in underestimates of prevalence of sensitive issues such as FGM, domestic violence or drug use. In a more detailed review, Alkire and Samman (2014) conclude that it is problematic to use perceptual data as proxy for objective deprivations.

Perception questions, as defined in this report, are not concerned with the individual *experience* or *condition* but rather with the *perception*. As such, framing questions in such a way that they do not imply direct disclosure of a socially undesirable behaviour may reduce the reliability bias, although it may not completely eliminate it. For example, in a study on perceptions of FGM in Ethiopia, de Cao and Lutz (2014) use a list experiment[36] as an alternative for direct questioning to obtain perceptions about FGM. This indirect method resulted in higher disclosure of acceptance of the practice, compared to direct questioning. This method has also been used to ask about racial attitudes, voter behaviour, the social acceptability of sexual attitudes and illegal migration.

The potential to reduce bias by using indirect questioning is promising, but a deeper issue is that perceptions fluctuate across groups of society and over time, and responses show systematic differences. These fluctuations need to be separated from reliability issues.[37] Distinguishing between true changes in perceptions, variability across types of respondents or some type of bias remains a contested issue, and it is as yet unclear what the goalpost should be for correcting the potential biases. For example, in the study of FGM attitudes in Ethiopia for example, the authors attempt to find biases by comparing the answers from direct and indirect questions; finding that uneducated women who had previous contact with a local NGO were more likely show a social desirability misreporting bias (de Cao and Lutz, 2014). In comparison, Ivanyna and Shah (2009) compare country rankings of state capacity using citizen or expert perceptions. They adjust the citizen-based answers for the possibility that respondents are afraid to tell the truth about their governments (measured by the freedom of expression in their country), are indoctrinated (when respondents think well of their governments because of exposure to favourable mass-media coverage) or are excessively optimistic or critical. These corrections are very interesting, yet depart from the premise that such biases are driving the differences between the expert-based and citizen-based perceptions of the state. In this case, the existence of reliability bias may be harder to prove than to correct.

c) Adaptation: have poor people come to terms with living in poverty?

Another aspect that affects the reliability of perception data is adaptation, particularly relevant when considering that, unlike the MDG, many if not all of the SDG targets are likely to be applied to all types of countries and peoples – poor and rich, and from a variety of backgrounds and cultures. Unlike the reliability issues, the adaptation problem does not imply that people misreport answers. In fact, people report their actual perception, but the interpretation of results becomes difficult. While this does not undermine the value of perceptions data, it affects how they should be used by policy-makers – for example, to understand how to approach people with ideas and potential policies to make improvements to their lives, rather than to decide whether or not to make such improvements.

There are two types of adaptation. The first results from habituation to one's living conditions and the consequent abandonment of wants, aspirations and expectations (Neff, 2012). If adaptation has taken place, the assessment individuals make of their wellbeing (or other aspects of their lives) may be limited by their life circumstances. Sen (1999) for example, argued that:

36 This is an indirect way to ask a question. The method presents respondents with a list of statements and asks them indicate the total number of statements with which they agree. The respondents are randomly divided into a control and a treatment group. The control-group respondents receive a list of non-sensitive items. The treatment-group respondents receive the same list of non-sensitive items plus one sensitive item, for example about FGM (e.g., 'A girl should be circumcised'). The difference in the total number of items between control and treatment group identifies the proportion of people in the population that agrees with the sensitive item.

37 For standard surveys, a common way to assess reliability is to apply questionnaires various several times to similar subjects and in similar conditions and compare the responses. Reliable questions would result in the same answers reported each time the questionnaire is applied.

'The deprived people tend to come in terms with their deprivation because of their sheer necessity of survival, and they may, as a result, lack the courage to desire any radical change, and may even adjust their desires and expectations to what they unambitiously see as feasible.' (Sen, 1999: 63)

Adaptation may change the place in the reference scale where individuals place themselves when answering perception questions. For example, in the questions about income sufficiency presented in Box 1, what constitutes 'sufficient' may be determined by one's actual living conditions. In this scenario, people living in poverty may say that their income is sufficient, while someone with higher incomes may say it is not. In consequence, if such adaptation exists, it limits the comparability of the answers for these two types of persons. This is also problematic in terms of comparability across groups because it makes it hard to determine whether the reasons behind the variability of responses are true differences in perceptions or are the product of adaptation.

The second type of adaptation occurs when extraordinary events temporarily change people's perceptions and bias subjective responses, particularly around subjective wellbeing. However, research has shown that such biases tend to be only short-lived and subjective wellbeing returns to a stable level after the events. This is called hedonic adaptation. For example, the change in reported satisfaction that happens after important life events such as getting married, having children or winning the lottery is often temporary and the overall effect on life satisfaction or long-term happiness measures is minor (Kahneman and Krueger, 2006; Gilbert, 2004). However, the evidence seems mixed on the degree of hedonic adaptation. Some have found that people adapt more to positive than to negative shocks (Burchardt, 2005); others have found that for health shocks such as disability onset, full hedonic adaptation does not take place (Fafchamps and Kebede, 2007; Oswald and Powdthavee, 2008). Large and randomly chosen samples would help to control for unobserved effects when those are unsystematically present (Arnesen and Trommald, 2005); using a sample drawn from a group with enough variation, one can expect that, in aggregate, indicators would be more reliable (Rodriguez Takeuchi, 2014).

It is important to highlight that empirical evidence suggests that the issue of adaptation, at least in the case of income poverty and subjective wellbeing, is less strong than initially supposed (Box 2). In fact, recent research shows that deprivation does not need to translate into adaptation. Findings suggest that subjective wellbeing is dependent on the fulfilment of basic needs (Veenhoven, 1988), but that income and basic needs fulfilment are not the sole determinants of higher subjective wellbeing. For example, Burchardt (2005) finds evidence of adaptation among the top income group only, suggesting that those on constant high incomes have adapted to that level. Similarly, the famous Easterlin paradox (1974) also shows that while country income has a positive correlation with subjective wellbeing, there seems to be a threshold above which the returns from income are lower. This was corroborated by Camfield et al. (2009) at the micro level.

Adaptation can be seen in many other areas. For example, research on women's demands for gender equality show that in gender-unequal countries (measured by objective indicators), the demand for gender equality is stronger (Bhatkal, 2014). Paradoxically, while these women demand equality, they also continue to hold views which reinforce gender inequality. For example, with respect to the right to education and work they may still express agreement with the statement that boys have more right to education than girls. This hints at the presence of adaptation because it suggests that women in more unequal societies tend to undervalue women's rights. In this sense, peoples' frames of reference are affected not just by their own (absolute) living conditions, but also by comparison to those around them. This situation may still point to the presence of very strong norms on accepted social attitudes, and thus would remain a good indicator of the evolution (or lack of it) of informal gender norms in a given country, as discussed in Section 3a.

However, if the adaptation bias is systematically present for one subgroup (i.e. women, the poor or those more exposed to mass media, as in the case presented by Ivanya and Shah (2009), this could be corrected for using econometric techniques, for example (see Decancq et al., 2009). As discussed, the main issue is that identifying the presence of a bias is not straightforward, as there is no gold standard or clear goalpost of comparison. Moreover, discarding systematic variations in responses across groups is not desirable. Rather, one advantage identified for perceptions data is to highlight those differences, as well as the opinions of groups that otherwise could be made invisible by using averages. Cross-validation of indicators with other methodologies is always desirable, and necessary when using any kind of indicators for national policy diagnosis.

d) Objective and subjective mismatch

In line with Stiglitz et al. (2009), we argue that objective and subjective indicators could be complementary rather than alternatives, as they might be measuring different aspects of a phenomenon. In that case, there is no reason to expect perfect coincidence between the two types of data. Rather than being a problem with perceptions data, a mismatch between the two types shows that they each contribute different information that is useful and necessary. If the two were giving the same information,

Box 2: Adaptation, income poverty and subjective wellbeing

Some research on subjective wellbeing has found a possible impact of adaptation in the interpersonal comparison of subjective wellbeing responses. In particular, the concern relates to poorer people reporting high levels of subjective wellbeing, despite grim living conditions. Contrary to popular belief, there is no widespread evidence of adaptation among the poorest (Clark, 2012). For example:

- Neff (2012) shows that in two villages in Andhra Pradesh, the majority of the poor are less satisfied with their lives on average compared to their non-poor counterparts, as are lower castes and Muslims compared to higher castes. He also finds highly significant statistical differences regarding the mean satisfaction level between expenditure quintile groups. The lowest expenditure quintile group has the lowest mean satisfaction.
- In Bangladesh, Camfield et al. (2009) find that, although the majority of people are satisfied with their lives, there are significant differences between richer and poorer groups. Of the rich, 31.5% of people report being 'very happy', compared to 4.7% of the poor.
- Mexico's first national subjective wellbeing survey (BIARE), conducted in 2013, shows that individuals from the lowest income decile are less satisfied with their lives than individuals from the highest income quintiles (Chávez, forthcoming).
- In South Africa, a study on social perceptions of material necessities (Wright, 2008) found that people lacking essential items defined by the study (such as having a bath or shower in the house) are aware of their shortages and listed them as essential needs. The study also finds consensus between different groups about what they define as essential needs and what a standard of living comprises.

These findings in India, Bangladesh, Mexico and South Africa indicate that people in lowest income groups report lower levels of subjective wellbeing and life satisfaction than those in higher income groups. At the country level, this also seems to hold. The Legatum Commission (O'Donnell et al., 2014) analysed GWP data and find that more developed Nordic countries regularly lead the world in terms of life evaluation or satisfaction, while less economically developed countries such as Togo, Sierra Leone, and Zimbabwe, are consistently at the bottom.

there would not be a need for using perceptions data, but rather efforts would need to be focused on perfecting the existing indicators and methods of data collection.

Stiglitz et al. (2009) point out that a reason behind the mismatch between subjective wellbeing and traditional economic indicators of country progress (i.e. GDP) is the incompleteness of the latter in terms of capturing some phenomena which have a large impact on wellbeing. While GDP per capita may be a good indicator of economic activity, it is insufficient to measure a country's progress in wellbeing, a concept that encompasses much more than national economic progress. Moreover, the relationship between objective and subjective wellbeing may be non-linear, that is, it changes depending on the specific levels of the indicators. As discussed in the context of adaptation and SWB measures (see Box 2), material living conditions, including income, are important determinants of subjective wellbeing at low levels. But once the most basic material needs are covered, the relationship may become more indirect, and other aspects –such as mental health, community and family relationships or environmental conditions – can play a greater role in driving wellbeing.

The mismatch may also arise in other types of indicators. Headey and Ecker (2013) find a very low correlation between self-reported and other types of measures of food security. This is partly because the different indicators are measuring different aspects of food security. For instance, the 2010 FAO State of Food Insecurity report used data mainly on food availability and average caloric requirements. This placed a large emphasis on food importing, and underestimated the role of national food production and distribution within countries. The food insecurity experience scale for the Voices of the Hungry project[38] does not provide specific information on actual food availability, but is useful to measure the psychological dimension of food security (Headey and Ecker, 2013), focusing on the socio-cultural aspects of the experience of hunger (Ballard et al., 2013).[39] Nevertheless, because of the issues around adaptation and cross-cultural understandings of hunger and food adequacy, these types of indicators may present problems in the identification of needs, for example to target programmes of food relief or food transfer.

Another interesting case is the measurement of governance. Foresti et al. (2014) and Bergh et al. (2014)

38 http://www.fao.org/fileadmin/templates/ess/voh/FIES_062014.pdf

39 There is also an issue around capturing inequality. While hunger is more an issue of access than of availability, food availability and caloric consumption can indicate only the upper bound of chronic nutrition in a country (for example, the situation where all food available in the country was equally distributed, and all individuals had similar caloric requirements).

have pointed to the difference between indicators looking at *forms* of governance versus indicators looking at *ways* of governance, a discussion that has been prominent in the post-2015 governance debates. For example, indicators measuring regime type or the existence of certain legislation or institutions (e.g. an anticorruption commission) may be prone to 'isomorphic mimicry' or the adoption of *forms* to meet the targets without real change in performance (Foresti et al., 2014). Thus, the gap between perceptions and objective indictors of governance may suggest a gap between an ideal form and the functioning of governance institutions and can signal perceived citizen dissatisfactions and grievances. Bergh et al. (2014) argue for example that citizen dissatisfaction with the lack of perceived responsiveness and functioning of democratic governments may have been behind many of the recent popular uprisings around the world.

In the area of governance, population surveys have the advantage of capturing the views of those directly involved in the institutions of the country, but could be less apt for cross-country comparisons than other sources of data such as experts' perceptions or objective data (Hulme et al., 2014). However, these authors argue that perceptions data are also less suitable for distinguishing between the different attributes of governance and thus harder to use for identifying relevant policy interventions. For example, it would not be clear what is behind a change in reported dissatisfaction with the police: an increase in crime, or an increase in police effectiveness in capturing criminals and thus the visibility of crime. This suggests that perception indicators of governance would not meet Foresti et al.'s criterion of showing a clear 'theory of change', that is, a 'clear pathway of change underpinning the proposed indicator, specifically considering how in practice it could lead to better monitoring and measurement, and hence greater action and resource mobilisation to improve governance outcomes' (Foresti et al., 2014: 18).

To summarise, in this section we have showed that, while both types of data often suffer from similar methodological difficulties related to the design and execution of surveys, perceptions data have additional problems which still require more research. In particular, in terms of ensuring the reliability of the data and the indicators and ensuring cross-country and cross-group comparability, there is not yet a clear consensus on how to identify and control for potential biases.

Finally, more research is also needed to clarify the source of differences between objective and perceptions indicators and what the gaps between the two can tell us. They may both capture different aspects of a single phenomenon (Stiglitz et al., 2009) – for example, food security or governance – and thus both are still useful, depending on the purposes of measurement. In particular, perception indicators appear useful in identifying gaps and critical points for intervention, for example before malnutrition becomes manifest or when citizen dissatisfaction signals a failure of democratic governance. However, these gaps highlight the difficulties of eliciting the appropriate theory of change behind the indicators. Determining how to react to the information presented by perceptions data, or identifying the exact policy interventions behind a change in perceptions, is a harder task. This challenge also occurs with other types of data and methods; mixed-methods approaches have emerged as an attempt to capture the complexity of how policies and programmes operate.[40]

40 See http://betterevaluation.org/resources/guides/intro_mixed-methods_impact-evaluation and http://betterevaluation.org/blog/mixed_methods_part1

Perceptions data in monitoring the post-2015 agenda: implementation issues

How can perceptions data be incorporated in the monitoring of key issues in the post-2015 agenda? This section outlines some practical issues to consider for implementation.

The first issue to consider is whether, in the context of the post-2015 targets, perceptions data would be more **useful** for national or international monitoring and tracking of the targets. While perceptions indicators could be potentially good markers of progress in certain areas, there are limits to their comparability, especially when considering responses given by people living in countries at very different levels of development (e.g. because of the adaptation problem). International rankings have emerged in numerous areas and are becoming an increasingly accepted policy tool in the international arena, but national reactions to rankings are not yet widely understood (Cooley, forthcoming).

Rankings inherently compare one country with others, and thus can confer status to countries ranked highly and stigmatise those ranked lower (Cooley, forthcoming). While the MDG and their SDG successors do not establish explicit country rankings, they do involve a large degree of comparison of progress across countries. Moreover, while the measurement and comparison of countries may have incentivised countries trying to reach the MDG targets (UNDP, 2010), it has been shown that this measurement was unfair, especially for the poorest countries for which some goals were unattainable (e.g. Easterly, 2009). The lesson here is that taking into account countries' heterogeneity and starting positions was important. Ranking or comparing countries based on their performance on perception indicators would likely be problematic for the post-2015 targets. Moreover, as discussed in previous sections, one of the great advantages of perceptions data are their potential to monitor changing conditions and to act as pointers of changes in trends over time. Because of the difficulty of discerning what is behind a change in a perception indicator, triangulation with other sources of data is necessary to determine the appropriate policy response. Thus, we suggest that perceptions data would be more useful to monitor changing situations over time within countries, rather than to establish comparisons across them.

A second, related issue is the **production** of perceptions data. Despite the increase in availability of perceptions data, few NSOs were convinced until recently of their potential to inform policy and thus few of them considered it worth investing in measuring such indicators. Thus, particularly in developing countries, surveys of population attitudes, expectations and satisfaction are not yet well established (UN Task Team, 2013).

On the other hand, private companies using perceptions data for market research have long realised their potential. Furthermore, private polling companies such as Gallup and Ipsos Mori are still among the leading collectors of subjective wellbeing data. Research consortia such as the European Value Surveys and the World Values Surveys are also leading on the collection of perceptions data. The fact that, currently, perceptions data are strongly associated with these institutions poses additional questions of data governance and accountability. The use of private sources as official measures for target-setting raises questions of data quality, ownership and accountability.

Data quality may not depend on whether the organisation is public, private or research-based. In fact, when private companies are being commissioned or work in partnership with public institutions, including international ones, to collect perceptions data, the boundaries of data ownership are hard to draw. Furthermore, in the case of research consortia, the presence of academics with strong research profiles may help to improve the quality of the data collected. However, it may be more difficult to assess the quality of the data if they are not subject to appropriate **accountability**.

For example, Gallup World Poll data have been used in international projects such as the World Bank Global

Findex,[41] UNDP's Human Development Report since 2010,[42] the OECD's Better Life index,[43] ILO's Social Unrest index[44] and World of Work Report, 2013 and the Legatum Prosperity Index.[45]

While some data are made public when they are part of such projects, the micro data behind the aggregate numbers are not released, and sometimes not even to a partner institution. Conducting further analysis, verifying the sampling procedures and representativeness of the surveys, checking for omissions, recording mistakes and other data quality issues is not possible without access to the micro data. It is also not possible to disaggregate the data beyond what the partner institution presents, for example, if they release results at only the national level. While the partner institutions could ensure data quality and accountability, this may not always be the case. Beyond that, this also implies a change in the way data are traditionally reported to track progress in the MDG. It will not be national governments reporting to the UN-system institutions, but such organisations collecting data from private organisations and then reporting back to countries on their progress.

Experiences such as that of the Philippines SWS demonstrate that there is value in private institutions generating social statistics for public use (Guerrero and Mangahas,[46] 1989, cited in Guerrero, 2004), especially in a context of a restricted democracy where governments may try to conceal information. The SWS was banned in 1982 by the authoritarian ruling regime for being 'too risky' to publish only when the SWS became independent; the results were widely disseminated across the general population (Guerrero, 2004). Counter experiences can also be found, especially in the case of political or marketing polls commissioned by interest groups to change public opinion, camouflage controversial policies, or gather 'proof' to sustain certain demands or concessions from the government (Dionne and Mann, 2003).

An increasing number of NSOs are starting to collect and release perception-based data, particularly on subjective wellbeing. To the extent that these data becomes useful, there is certainly consensus that SWB measures should be viewed as one set in the much broader array of indicators through which populations are monitored and policies informed (see OECD, 2013). The key issue seems to be whether the data are made available and communicated widely, rather than whether they are public or private. Availability also enables the improvement of data quality, as checks can be conducted by data users, and the strengths and limits of particular surveys can be better understood.

Further, as the OECD has done for indicators of SWB, a common set of standards will need to be derived for other perception-based indicators. While a wider range of perceptions data is collected on gender norms, violence and security, and governance, different surveys often ask different questions and a compilation and review of best practices is necessary. The review presented here, on indicators of gender social norms, violence and government accountability should serve as a starting point. Projects focused on perceptions data are also growing and increasingly looking to standardise their procedures. The WVS, and its precursor, the European Value Surveys, as well as the regional barometer network are examples where joint efforts have increasingly led to a common set of procedures. However, questions remain about the value added of a global perception survey that is fully comparable across countries, especially given that, for the SDG, perceptions data would be most useful at the national level.

While some consensus is starting to emerge on common standards for data quality and for collecting and analysing perceptions data, some degree of flexibility in how questions are asked is needed to be able to account for the country-specificity of perceptions. For example, to ensure that the framing of question does not result in biased estimates for sensitive issues, the specific techniques and terms used for the questions may need to be adapted depending on the country, as a simple translating of the questionnaires may not be sufficient.

Finally, it is worth noting that the application of standards of good-practice procedures, as well as the desired increase in the sample size to try to address representativeness and inclusion of marginalised groups, may also increase the cost of perception surveys. This may risk one of their key selling points – frequency and country coverage. Although a detailed review is beyond the scope of this paper, exploring the mix of new methods and new technologies in the collection of perceptions data could be a way to bring those costs down. More exploration of this aspect is required.

41 World Bank article: http://bit.ly/OGopsB

42 Although perceptions data are not included in the HDI

43 http://www.oecdbetterlifeindex.org/

44 http://www.ilo.org/newyork/voices-at-work/WCMS_217280/lang--en/index.htm

45 http://www.prosperity.com/

46 Note that Mangahas is the founder of the SWS.

Conclusions

The post-2015 negotiations and in particular the call for a 'data revolution' have spurred debate around the inclusion of new indicators of progress. In this context, indicators based on perceptions data could shift the measurement debate and help to focus on what is important to improve wellbeing of citizens around the world, in both objective and subjective dimensions.

The main advantage of perceptions data is that they can present the respondents' own views directly. This is important to stimulate public debate around the trade-offs of competing policies and to hold policy-makers to account for what matters most to people. To select targets and indicators to measure the post-2015 goals, this process can help to voice to citizen's perspectives on what they perceive as the constituent elements of the goals. Moreover, some perceptions surveys such as MyWorld can be used to prioritise implementation among the plethora of goals and targets and to compare the preferences of people in different countries or belonging to specific social groups.

The increase in the availability of perceptions data and the research around subjective dimensions of wellbeing present a great opportunity. In particular, compared to household surveys and other sources of information such as censuses, perception surveys have great potential to produce timely and frequent data for monitoring social conditions, and to spot changes in trends that could endanger the achievement of the post-2015 targets.

We have identified three areas where perceptions data have featured strongly in proposals for monitoring the SDG: gender and social norms, violence and security, and governance. In each case, perceptions data could provide complementary information in conjunction with other data. While perception surveys could be difficult to use on their own to measure the achievement of a target in any of these three areas, different considerations need to be incorporated in the decision-making process with perceptions information as a consistent input. Moreover, a disaggregated analysis of perceptions may help to make explicit the variety of preferences in a given society, which needs to be understood in the context of the post-2015 targets to address inequalities in order to 'leave no one behind'.

However, perceptions data are less useful in determining the appropriate policy change required when such an alarm is raised. The main limitations are: 1) representativeness at the sub-national level, particularly given low sample sizes; 2) reliability of the answers; and 3) the issues around adaptation and how to separate this from heterogeneity in points of view. These challenges are still unresolved, and may limit the comparability of answers, especially of those living in extreme deprivation or with restricted freedoms and rights.

References

Alkire, S. and Samman, E. (2014) *Mobilizing the household data required to progress toward the SDGs*. Sustainable Development Solutions Network. (http://bit.ly/1DpxLQq)

Alkire, S., Malapit, H., Meinzen-Dick, R., Peterman, A., Quisumbing, A., Seymour, G. and A. Vaz. (2013) 'Instructional guide on the Women's Empowerment in Agriculture Index'.

Arnesen, T. and Trommald, M. (2005) 'Are QALYs based on time trade-off comparable? A systematic review of TTO methodologies', *Health Economics* 14(1): 39-53.

Ballard, T., Kepple, A. and Cafiero, C. (2013) The Food Insecurity Experience Scale. Development of a global standard for monitoring hunger worldwide. Rome: FAO.

Bardasi, E., Beegle, K., Dillon, A. and Serneels, P. (2011) 'Do labor statistics depend on how and to whom the questions are asked? Results from a survey experiment in Tanzania', *World Bank Economic Review* 25(3): 418-447.

Barker, C., Pistrang, N. and Elliott, R. (2002) 'Self-reported methods', in *Research methods in clinical psychology. An introduction for students and practitioners* (2nd edn). Wiley.

Bartels, L.M. (2003) 'Is "popular rule" possible? Polls, political psychology, and democracy', *The Brookings Review*, Summer.

Bellani, L., Hunter, G. and Anand, P. (2013) Multidimensional welfare: do groups vary in their priorities and behaviours? *Fiscal Studies* 34(3): 333-354.

Bergh, G., Rocha Menocal, A. and Rodriguez Takeuchi, L. (2014) *What's behind the demand for governance? An assessment of people's views*. London: ODI.

Bhatkal, T. (2014) What do women want? Gender, perceptions data and development priorities. London: ODI.

Blanchflower, D. and Oswald, A. (2011), 'International happiness'. NBER Working Paper 16668. National Bureau of Economic Research.

Boarini, R., Comola, M., Smith, C., Manchin, R. and De Keulenaer, F. (2012) What makes for a better life? The determinants of subjective well-being in OECD countries: evidence from the Gallup World Poll. Paris: OECD.

Burchardt, T. (2005) 'Are one man's rags another man's riches? Identifying adaptive expectations using panel data', *Social Indicators Research* 74: 57-102.

Camfield, L., Choudhury, K. and Devine, J. (2009) 'Well-being, happiness and why relationships matter', *Journal of Happiness Studies* 10: 71-91.

de Cao, E. and Lutz, C. (2014) *Sensitive survey questions: Measuring attitudes regarding female circumcision through a list experiment* (http://bit.ly/145E2S5).

Carr-Hill, R. (2013) 'Missing millions and measuring development progress', *World Development* 46: 30-44.

Chandy, L. (2013) 'Counting the poor. Methods, problems and solutions behind the $1.25 a day global poverty estimates'. Investments to End Poverty Working Paper. Development Initiatives.

Chávez, C. (forthcoming) 'Youth Delinquency in Mexico: An exploration of contributing factors'. PhD diss., University of Cambridge.

Clark, D. (2012) 'Adaptation and development – issues, evidence and policy relevance', in D. Clark (ed.), *Adaptation, poverty and development: the dynamics of subjective well-being*. London: Palgrave Macmillan.

Clemens, M., Kenny, C. and Moss, T. (2007) 'The trouble with the MDG: confronting expectations of aid and development success', *World Development* 35(5): 735-751.

Cooley, A. (forthcoming) 'The emerging politics of international rankings and ratings: a framework for analysis', in A. Cooley and J. Syder (eds), *Ranking the world: the politics of international rankings and ratings*. Cambridge University Press.

Coulthard, S. Camfield, L. and McGregor, A. (2014) 'Valuing development and social progress: practical approaches to measure what matters in development'. London: ODI (mimeo).

Cummings, R. (2009) 'Measuring population happiness to inform public policy'. The 3rd OECD World Forum on 'Statistics, Knowledge and Policy' Charting Progress, Building Visions, Improving Life, Busan, Korea, 27–30 October 2009.

Decancq, K., Fleurbaey, M. and Schokkaert, E. (2009) 'What good is happiness?' Ecore discussion paper (http://www.ecore.be/DPs/dp_1238577455.pdf).

Dionne, E.J. and Mann, T.E. (2003) 'Polling & public opinion: the good, the bad, and the ugly', *The Brookings Review*.

O'Donnell, G., Deaton, A., Durand, M., Halpern, D. and Layard, R. (2014) *Report of the Commission on Wellbeing and Policy*. London: Legatum Institute.

Easterlin, R. (1974) 'Does economic growth improve the human lot? Some empirical evidence', in Paul A. David and Melvin W. Reder (eds), *Nations and households in economic growth: essays in honor of Moses Abramovitz*. New York: Academic Press.

Easterly, W. (2009) 'How the Millennium Development Goals are unfair to Africa', *World Development* 37(1): 26-35.

Fafchamps, M. and Kebede, B. (2008) *'Subjective wellbeing, disability and adaptation: A case study from rural Ethiopia'*. Working Paper 283. Oxford: The Centre for the Study of African Economies (http://www.bepress.com/csae/paper283).

Foresti, M. Wild, L. with Norton, A. and Rodriguez Takeuchi, L. (2014) *Governance targets and indicators for post 2015. An initial assessment*. London: ODI.
Gallup (2012) Global states of mind. New metrics for world leaders. Washington D.C.: Gallup.
Garrett, J. and Ahmed, A. (2004) 'Incorporating crime in household surveys: a research note', *Environment and Urbanization* 16: 139-151.
Gilbert, D. (2004) 'The surprising science of happiness'. TED Talks (http://www.ted.com/talks/dan_gilbert_asks_why_are_we_happy.html).
Gough, I. and McGregor, J.A. (2009) 'After 2015: "3D human wellbeing".' IDS in Focus, Policy Brief Issue 09 (http://www.ids.ac.uk/files/dmfile/IF9.2.pdf).
Groves, R.M. et al. (2009) *Survey methodology* (2nd edn). New Jersey: Wiley.
Guerrero, L.L.B. (2004) *Social weather stations: Asia's oldest barometer*. WSW Occasional Paper (June).
Harper, C., Nowacka, K., Alder, H. and Ferrant, G. (2014) Measuring women's empowerment and social transformation in the post-2015 agenda. London: ODI.
Harper, C., Jones, N. and Watson, C. (2012) Gender justice for adolescent girls: tackling discriminatory social norms. Towards a conceptual framework. London: ODI.
Headey, D. and Ecker, O. (2013) 'Rethinking the measurement of food security: from first principles to best practice', Food Security 5(3): 327-343.
Herbert, S. (2013) Perception surveys in fragile and conflict-affected states (GSDRC Helpdesk Research Report 910). Birmingham, UK: GSDRC, University of Birmingham (http://www.gsdrc.org/docs/open/HDQ910.pdf).
HLP (High Level Panel of Eminent Persons on the Post-2015 Development Agenda) (2013) A new global partnership: eradicate poverty and transform economies through sustainable development. New York: UN.
Hulme, D., Savoia, A. and Sen, K. (2014) Governance as a global development goal? Setting, measuring and monitoring the Post-2015 Development Agenda. ESID Working Paper 32. Manchester: Effective States and Inclusive Development (ESID).
Hussein, E. (2010) Women's experiences, perceptions and attitudes of female genital mutilation. The Bristol PEER Study. Foundation for Women's Health Research and Development (FORWARD).
IEAG (Independent Expert Advisory Group on a Data Revolution for Sustainable Development) (2014) A world that counts (http://bit.ly/1wApBjP).
Ipsos Mori (2012) Public perceptions of the NHS and social care: an ongoing tracking study conducted for the Department of Health, December 2011 Wave. London: Ipsos Mori.
Ivanyna, M. and Shah, A. (2009) Citizen-centric governance indicators: measuring and monitoring governance by listening to the people and not the interest groups. Economics Discussion Paper 2009-27, Kiel Institute for the World Economy (http://bit.ly/1wwvGLh).
Jerven, M. (2103) Poor numbers: how we are misled by African development statistics and what to do about it. Ithaca and London: Cornell University Press.
Kahneman, D. and Krueger, A.B. (2006) 'Developments in the measurement of subjective well-being', Journal of Economic Perspectives 20(1): 3-24.
Kaufmann, D., Kraay, A. and Mastruzzi, M. (2009) Governance matters VIII aggregate and individual governance indicators 1996-2008. Policy Research Working Paper 4978. Washington DC: The World Bank.
Kenny, C. and Dykstra, S. (2013) The Global Partnership for Development: a review of MDG 8 and proposals for the post-2015 development agenda. Washington DC: CDG.
Kroll, C. (2013) Global development and happiness: how can data on subjective wellbeing inform development theory and practice? IDS working paper 432. Brighton: Institute of Development Studies.
Krosnik, J.A. and Presser, S. (2009) 'Question and questionnaire design', in J.D. Wright and P.V. Marsden (eds), Handbook of survey research (2nd edn). San Diego, CA: Elsevier.
Leo, B. (2013) Is anyone listening? Does US foreign assistance target people's top priorities? CDG Working Paper348 (http://bit.ly/19KhMtx).
Lucci, P. and Bhatkal, T. (2014) Monitoring progress on urban poverty: are current data and indicators fit for purpose? London: ODI.
Mallett, R. (2012) How far do perception surveys take us in fragile and conflict-affected situations? The Asia Foundation (http://bit.ly/1tQGAA0), accessed 28 November 2014.
Mangahas, M. (2013) The value of statistical competition. 12th National Convention on Statistics, Mandaluyong City, October 1-2, 2013.
McGee, R. and Gaventa, J. (2011) Synthesis report: review of impact and effectiveness of transparency and accountability initiatives. London: Transparency and Accountability Initiative.
Melamed, C. (2012) Does development give poor people what they want? London: ODI.
Møller, V. (2005) 'Resilient or resigned? Criminal victimisation and quality of life in South Africa', Social Indicators Research 72(3): 263-317. Kluwer Academic Publishers.

Neff, D. (2012) 'Adaptation, poverty and subjective well-being: evidence from South India', in D. Clark (ed.), Adaptation, poverty and development: the dynamics of subjective well-being. London: Palgrave Macmillan.
OECD (2013) OECD guidelines on measuring subjective well-being. Paris: OECD.
Oswald, A.J. and Powdthavee, N. (2008) 'Does happiness adapt? A longitudinal study of disability with implications for economists and judges', Journal of Public Economics 92(5-6): 1061-1077.
Prizzon, A. (2014) What do countries want from aid? Some reflections from the field. Valuing Progress blog series (http://bit.ly/1HUa1nj).
Riphahn, R.T. and Serfling, O. (2002) Item non-response on income and wealth questions. IZA discussion paper N576. Bonn: IZA.
Rodriguez Takeuchi, L. (2014) Incorporating people's values in development: weighting alternatives. London: ODI.
Saferworld (2013) Addressing conflict and violence from 2015: a vision of goals, targets and indicators. February 2013 Briefing. London: Saferworld.
Samman, E. and Rodriguez Takeuchi, L. (2014) Old age, disability and mental health: data issues for a post-2015 framework. London: ODI.
SciDev (2014) 'Frustration' over old or missing data in MDG report, 8 July 2014 (http://bit.ly/1xISWtj).
Seaford, C. (2013) 'The multiple uses of subjective wellbeing indicators', Social Indicators Research 114: 29-43.
Sen, A. (1999) Development as freedom. Oxford: Oxford University Press.
Stiglitz, J.E., Sen, A. and Fitoussi, J.-P. (2009) Report by the Commission on the Measurement of Economic Performance and Social Progress (http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf).
Thomas, M.A. (2009) 'What do the Worldwide Governance Indicators measure?', European Journal of Development Research, 16 July.
Tortora, R.D. (2009) Sub-Saharan Africans rank the Millennium Development Goals (MDG). Washington DC: Gallup.
Tourangeau, R. and Yan, T. (2007) 'Sensitive questions in surveyss', Psychological Bulletin 133(5): 859-883.
Tourangeau, R., Groves, R. and Redline, C. (2010) 'Sensitive topics and reluctant respondents. Demonstrating a link between nonresponse bias and measurement error', Public Opinion Quarterly 74(3): 413-432.
Traugott, M.W. (2003) 'Can we trust the polls? It all depends', The Brookings Review, Summer.
UN (2014a) Introduction to the Proposal of the Open Working Group for Sustainable Development Goals. New York: UN.
UN (2014b) The road to dignity by 2030: ending poverty, transforming all lives and protecting the planet. Synthesis Report of the Secretary-General on the Post-2015 Agenda. New York: UN.
UNDP (2014) Building transparency, accountability and anti-corruption into the post-2015 development framework. UNDP Discussion Paper, September 2014. New York: UN.
UNDP (2010) Beyond the midpoint: achieving the Millennium Development Goals. New York: UNDP.
UN Task Team (on the Post-2015 Development Agenda) (2013) Statistics and indicators for the post-2015 development agenda. New York: UN (http://bit.ly/12xWbHF).UN Women (2013) A transformative stand-alone goal on achieving gender equality, women's rights and women's empowerment: imperatives and key components. New York: UN Women.
Veenhoven, R. (1988) 'The utility of happiness?', Social Indicators Research 20: 333-354.
Veenhoven, R. (2004) Subjective measures of well-being. UNU-WIDER Discussion Paper 2004/07 (http://bit.ly/1KoTkmh).
World Bank and SecondMuse (2014) Big data in action for development (http://bit.ly/1DEafwG).
Wright, E. (2008) Findings from the Indicators of Poverty and Social Exclusion project: a profile of poverty using the socially perceived necessities approach (http://bit.ly/1AKnmMY).
Yirga, W. et al. (2012) 'Female genital mutilation: prevalence, perceptions and effect on women's health in Kersa district of Ethiopia', International Journal of Women's Health 4: 45-54.

Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
Tel +44 (0)20 7922 0300
Fax +44 (0)20 7922 0399
www.odi.org
info@odi.org