

**Using Representative Opinion Surveys in the
African Peer Review Mechanism Process**

Robert Mattes

Associate Professor, Dept of Political Studies
Director, Democracy in Africa Research Unit and Centre for Social Science Research
University of Cape Town
rmattes@cssr.uct.ac.za

Deputy Director, Afrobarometer
(www.afrobarometer.org)

Presented to
APRM Lessons Learned – A Workshop for Practitioners, Researchers and Civil Society
Johannesburg, South Africa

12-13 September 2006

Using Representative Opinion Surveys in the African Peer Review Mechanism Process

Robert Mattes

Any process of national self-review would be incomplete if it only included the assessments of elites (whether they be government officials, technocratic experts, or civil society stakeholders) and excluded the opinions of the mass public. The true state of political and economic governance in a given country cannot be assessed simply on the basis of an objective analysis of the rules, resources and behavior of the economy, government institutions and large corporations.

Competent business people would never draw a final conclusion about the quality of their company and product simply on the basis of an investigation of the company charter, its internal processes or the assembly line. They would also need to know whether the consumers are actually buying their product, but more importantly whether or not they are satisfied with the product and likely to keep buying it or switch to some other product. Much in the same way, the actual state of political governance and, especially, democratic politics, is at least partially in the eye of the beholder.

But exactly *how* the values, awareness, evaluations and experiences of ordinary people are to be gathered is not a simple matter. On one hand, a country may wish to instill a sense of *public ownership* over the project and encourage the *participation* of as wide a cross section of ordinary citizens as possible. On the other hand, any self-assessment that aims to provide a true reflection of the actual state of affairs in the country would want to be as *accurate*, and therefore as *representative* as possible. The difficulty is that for a range of methodological, pragmatic, and socio-political reasons, it is rarely possible to maximize both these goals at the same time.

“Participatory” Consultations of Public Opinion

One apparent way to consult public opinion and simultaneously instill a sense of awareness and public ownership over the process is to run as broadly a consultative process as possible in which some sort of enumerators speak to ordinary citizens in their homes or in public meetings and record their responses, either through the forms of structured responses to structured questionnaires, or transcripts of semi structured or unstructured discussions and debates.

Public discussions have many advantages. Most importantly, they allow people *set the agenda, name their problems* and *frame the issues and range of potential solutions* in their own words rather than having them structured by the questionnaire designers. Moreover, they are deliberative, meaning that people can persuade each other to change their opinions through the course of the discussion.

However, public consultations also have many disadvantages. First of all, people are not political animals. The daily necessities of earning a living, family life, friends, and social activities compete with people’s attention to public affairs and their willingness to take part in political events. Thus, even the most well funded public consultation exercise may engage the attention of only a small fraction of ordinary citizens, let alone get them to participate – especially if people do not see any real incentive in doing so. The recent South African exercise is a classic case in point. The January-February 2006 South African Afrobarometer survey found that just one in twenty people (6 percent) said they had even heard about the APRM process, 1-in-33 (3 percent) attended a public meeting, and just 1-in-50 had filled out a questionnaire.¹

Second, virtually any process of public consultation means that citizens have to take the initiative to come out and make their voices heard. And if there is one thing that we have learned from fifty years of studying political behavior, it is that not all people are equally endowed with the same levels of willingness or capacity to take time out from their daily schedules and talk about politics. Thus,

¹ “Public Participation in South Africa’s African Peer Review Mechanism: Results from the January-February 2006 Afrobarometer – South Africa,” Presented to “APRM Lessons Learned – A Workshop for Practitioners, Researchers and Civil Society,” 12-13 September 2006, Johannesburg, South Africa.

consultative campaigns may not only fail to reach meaningful sizes of ordinary citizens, they may also fail to reach a representative cross section. Again, we can see clear evidence of this in the Afrobarometer examination of the South African APRM process. Not only were the better educated, frequent newspaper readers and active members of civil society organizations and trade unions more likely to have heard about the process, they were also more likely to have attended a meeting and filled out a questionnaire than the rest of the citizenry.² This then is inherently not a representative cross section of the public. People who have more accesses to education and the mass media, who are more attentive to politics, and who are more actively engaged in civil society, are likely to have significantly different values and opinions than citizens who are not.

Thus, the desire to consult a broad cross section of ordinary citizens, and allow them to participate in these important processes, and in turn instill a sense of national ownership of the process, can easily backfire, both failing to reach a significant section of the public and producing a potentially biased view of public opinion.

Representative and Accurate Assessments of Public Opinion

Ironically, the problem of goal of obtaining a representative assessment of public opinion is not solved by contacting ever larger numbers of citizens. Rather, the solution is in the *method* by which citizens are selected, much more than how many citizens are selected.

In other words, the solution is in sampling citizens, rather than attempting to create a mini census. The representativeness of a sample (the extent to which it produces estimates of public opinion or experience that mirror that of the total population) depends on two criteria. First of all, rather than allowing people to participate on their own initiative (which produces the well-known biases outlined above), the process of selecting individuals for consultation must be random. Second, every citizen of the country must have an equal and known chance (or probability) of being selected.

The *accuracy* of any estimate taken from a sample does, however, depend to some extent on numbers. A wealth of past experience shows us that even a random probability sample of 300 people can produce estimates about some population that are accurate (95 percent of the time) to within a margin of sampling error of plus or minus 5 percentage points. However, very few analysts would be satisfied with knowing, for example, that the fact that 45 percent of respondents say they are satisfied with the performance of the President means that Presidential approval in the total population lies somewhere between 40 and 50 percent.

While we can increase the accuracy of the results by spending more money and contacting more people, the law of diminishing returns that lies behind the mathematical basis of sampling means there is no 1-to-1 return. In order to reduce the margin of sampling error by one percentage point, we need to double the sample size.

<u>Sample Size</u>	<u>Sampling Error</u>
300	+/- 5 points
600	+/- 4 points
1,200	+/- 3 points
2,400	+/- 2 points
4,800	+/- 1 point

Thus, we soon reach a point of diminishing returns where the added costs of massively increasing sample sizes (say by contacting 4,800 respondents rather than 2,400) only brings marginal returns in increased accuracy. This is why the large majority of socio-political based surveys use sample sizes somewhere between 1,000 and 2,500. We are generally satisfied with knowing that satisfaction with Presidential Performance hovers somewhere between 44 and 48 percent. We do not get too concerned over whether 18 or 22 percent of all citizens actually contacted their MP in the previous year as long

² “Public Participation in South Africa’s APRM,” 2006.

as we can draw a broad inference that approximately one in five people did so. While we would like to be more accurate, its simply costs too much.

In contrast, national statistical agencies often run much larger household-based surveys because they place much greater emphasis on statistical precision on which to base development policy. It really does matter whether the actual rate of unemployment is 40 or 41 percent. Larger sample sizes also allow more accurate inferences from smaller subgroups in the sample to their counterparts in the population. Is there a difference, for example, in job seeking strategies between young, unmarried urban men, and young, unmarried urban women?

Thus, there is a bit of a paradox here. The attempt to consult larger and larger numbers of people and allow them to participate in the process of national self review may ultimately fail to contact a meaningfully large number of people, and more importantly, is almost guaranteed to produce a biased picture of public opinion. In contrast, surveys of relatively small but representative, random probability samples of ordinary citizens can produce accurate and cost effective estimates of public opinion.

Thus, surveys of random, probability samples of ordinary citizens are an absolutely essential part of the process of national self-review. It is true, however, that because opinion surveys are largely based on structured questionnaires, they allow the questionnaire designers to set the agenda, name the issues, and frame the allowed responses. But even with these drawbacks, they can be defended by the basic principles that seem to necessitate a consultative, participatory process. That is, representative surveys by their very nature, treat all citizens equally and offer everyone an equal and known chance of being selected to participate in, and thus influence the self review process (even if that participation is simply answering a set of questions).

Mechanics of Representative Surveys: Essentials

But besides saying that representative surveys are an irreplaceable element of the national self review process, there are a series of *essential elements* that have to be in place to ensure that the surveys are credible.

Absence of Widespread Civil Conflict

First of all, the freedom to travel and visit people virtually anywhere in the country is a prerequisite for fieldworkers to be able to contact a nationally representative sample. This means the absence of widespread civil conflict, politically hostile “no go” zones, crime or other obstacles such as natural disasters or large tracts of unmapped landmines that could compromise the safety of fieldworkers. But how much is too much? In general, there is no simple statistical answer to this question. The key question is the degree to which excluding these areas from the sample would compromise our ability to generalize from the responses of the rest of the country.

Accurate and Recent Census Data

Another prerequisite to conducting credible surveys based on representative samples is the availability of recent, accurate census data that is sufficiently detailed to allow disaggregation to quite small areas, even to the level of the basic census Enumerator Area. This is important because we begin the sampling process (a multi stage approach is discussed below) by disaggregating the census into a list of its smallest geographic units (e.g. Enumerator Areas) and then picking from this list a sample of these units. However, because these units often differ in size, we need to know the actual population size of each unit in order to weight its probability of selection. If each unit has an equal probability of selection regardless of its population size, the sample would no longer be representative. This principle of sampling is what we know as “probability proportionate to size” (PPPS).³

³ It should be noted that we may from time to time use disproportionate sampling for the purposes of providing reliable estimates of small, but socially or politically relevant groups so long as they are subsequently weighted back down to their true proportion of the population.

However, simply picking a sample of geographic units from one national list may randomly and unintentionally fail to include politically important areas or groups, or fail to reflect important variations across the population. Thus, the census data should also be sufficiently detailed to allow us to stratify, or cluster these units into a larger number of sub-lists that reflect politically relevant lines such as rural-urban differences, religious or linguistic differences, or districts and provinces. The principle of PPS also means that the census must tell us the relative population size of each of these strata, or sub-lists so that we do not select too many or too few units from each of these lists.

Finally, the census should also be sufficiently detailed to enable us to examine certain demographic data that can only be collected once we select and interview a respondent (e.g. age, marital status, income, education) and compare it to the actual population figures in order to assess the representativeness of the realized sample and decide whether it is necessary to weight the sample in order to make it represent the national demographics.

Multi Stage Sampling

As hinted at above, the sampling process is a multi-stage one. Few countries -- fortunately -- have a national list of all citizens, or at least one that they would share with a survey firm. So we have to sample citizens by first sampling the things in which we know they live -- households. But we also rarely have a single unified list of all households. Thus:

- Stage 1 consists of the process outlined in the previous section: that is, randomly sampling small geographic units from a national, stratified list of all those units based on the principle of PPS.⁴
- Stage 2 consists of sampling households within the selected geographic units.
- And Stage 3 consists of sampling individuals from within the selected households.

*The Importance of Stage 3: A Sample of Heads of Households is **Not** A Sample of Citizens*

The proper implementation of Stage 3 is absolutely imperative if we want to say that our survey results are representative of ordinary citizens. We want a sample of people, not a sample of households. As explained above, households are simply a convenient place to find people.

This means that we should guard against uncritically accepting the standard sampling procedures of national census or statistical institutions. The social issues that these institutions are generally interested in monitoring are traditionally addressed by economists and sociologists with household surveys because they have defined the household as a critical unit of analysis. The head of household is then conventionally selected to act as an informant about the status, activities and experiences of the household, and basic demographic data is collected for all individuals in the household.

But households have a series of properties that are important to economists, sociologists and development planners but which are simply not a factor for those interested in issues surrounding democratic citizenship. When it comes to democracy and governance, the individual citizen, not the household, is the proper unit of analysis.⁵ The very theory of citizenship on which democracy is

⁴ Some surveys may insert a prior stage in order to reduce travel costs in which they first create a list of larger geographical units, such as counties or districts, that are not too large, but relatively numerous and which group Enumerator Areas into fairly homogenous clusters. This list should also be stratified along rural-urban, or provincial lines. Once a small list of these larger clusters is drawn, a sample of Enumerator Areas can then be drawn from within the selected clusters.

⁵ Some of the early APRM national surveys (e.g. Kenya) appear to suffer from this problem, as well as apparently all the UNECA household surveys. A simple explanation is that those who designed these surveys had strong backgrounds in socio-economic household surveys and simply copied the sample design over to a public opinion survey without carefully thinking through the consequences.

premised stresses that all legal citizens should have as equal influence as possible on the affairs of government, including in the public input to the national self review.

But this is more than an issue of democratic ideology. As I shall discuss in greater detail below, any survey instrument that is specially designed to generate information to enable analysts to complete the self review questionnaire would typically ask about a wide a range of evaluations and preferences, as well as behaviors and knowledge. I can think of only a very small set of questions – usually relating to household finances – about which the head of household might have superior knowledge compared to other legal citizens who live in that household. But for the vast remainder of questionnaire items, any analyst attempting to fill out the self review questionnaire would have no reason to privilege the experiences, behaviors or opinions of the head of household over those of the rest of the citizenry.

Finally, this is also an issue of representativeness and accuracy. Because heads of household are more likely to be older, employed and male, and because they have more responsibilities that may lead them to look at the world differently than other household members, only interviewing heads of household is very likely to provide biased and misleading results.

Thus, the uncritical use of household survey methodologies in the national self review process may end up wasting huge amounts of money because the results will only be generalizable to heads of households, not to all citizens.

Accurate Translations

In order to be both representative, and to enable all citizens to have an equal influence on the overall results, it is imperative that all respondents are able to hear and respond in the language in which they feel most comfortable. Thus any survey instrument conducted in Africa should be translated – word for word, not just key concepts -- from English, French or Portuguese into all relevant home languages based on the drawn sample.

Minimum Sample Size

For the reasons discussed above, any survey claiming to be national in scope should interview at least 1,200 respondents, which would provides estimates of the national public that are accurate to within plus or minus 3 percentage points.

Mechanics of Representative Surveys: Desirables

There are a range of factors that should ideally be in place to carry out a credible and representative survey, but can be seen as “*desirables*” rather than “*essentials*.”

Household Lists and Maps

The “international gold standard” for survey research requires that samples be selected using PPS at all stages of sampling.⁶ As discussed above, any decent census should enable African survey researchers to select enumerator areas or other geographic units based on probability. But selecting households based on probability requires that we have an up to date list of all households in the selected sampling unit, and if possible, information about the size of each household. However, many African censuses cannot provide this detailed level of information, and if they can it is often hopelessly out of date. If the census office cannot provide this level of detail, there is the option of having the fieldworkers arrive in the sampling unit a day or too ahead of time and constructing the map themselves. However, this increases the accommodation and per diem component of fieldwork costs and drives up the overall cost of the survey.

⁶ Anthony Heath, Stephen Fisher and Shawna Smith, “The Globalization of Public Opinion Research,” *Annual Review of Political Science* 8 (2005): 297-333.

However, I do not regard such a level of precision as absolutely essential. If it is not possible or too costly to get this information, African survey researchers can and should at least satisfy a “silver standard”: that is, as long as all enumerator areas are chosen based on PPPS, it is reasonable to select households and respondents with random methods which are strictly monitored and enforced by field supervisors and over which the fieldworker has no control (such as randomized starting points in the EA, randomized walk paths stopping at every n^{th} house, and randomly varying this interval each day, and a random rule of selecting amongst eligible household members).⁷

Substitution

Again, the “international gold standard” holds that survey researchers should not allow any substitutions of selected respondents or households who refuse or are unable to be interviewed; this guards against ending up with a biased sample that under represents economically active people, or groups who do not feel comfortable talking about their social and political attitudes.⁸

If researchers are worried about large rates of non-response, they can either draw overly large samples ahead of time, or depending on the level of non-response draw new and separate smaller samples after the fact and interview the entire sample. But the first option presumes fairly sophisticated knowledge about past response rates that we rarely have due to the relative recency of survey research in most African countries. The second option often entails intolerably large increase in fieldwork costs.

But it is not clear whether allowing substitution necessarily results in major biases. Again, survey researchers in Africa may reasonably hold costs under control yet satisfy a “silver standard” if they only allow substitution of one household for another (but never substitution within households of one respondent for another), only allow it after at least two or three attempt to reach the targeted household and respondent, and only if they keep accurate data that would allow a post hoc comparison of the responses of substituted and non-substituted respondents.⁹

Survey Timing

Finally, it is desirable, though not essential, that survey designers are able to plan ahead sufficiently to allow them to conduct surveys supporting the national self review process in as politically neutral a period as possible. Essentially, this means not conducting surveys in the run up to or immediate aftermath of elections, and trying to avoid any other times in which the national mood might be artificially, but predictably optimistic or pessimistic.

Questionnaire Content

We need to approach the design of an APRM related public opinion questionnaire with a sensible theory of governance and democracy and the role of the citizens within this. It should begin with an examination of what the principle of fundamental equality and equal influence means for the content of a survey questionnaire aimed at citizens, yet this should be balanced with a keen sense of what citizens are and are not able to tell us.

For example, there are a whole range of political issues about which citizens have a right to express their preferences and evaluations regardless of whether or not they are based on any real experience or other information, for example, evaluations of elected leaders and most public institutions. In this area, the perception is a very large part of the reality that a national self assessment process needs to measure. Regardless of whether a given government department is actually a hotbed of nepotism, the popular perception that it *is* is probably more important than the actual state of affairs.

⁷ Robert Mattes, “Public Opinion in Emerging Democracies: Are the Processes Different?” *Handbook of Public Opinion Research*, Wolfgang Donsbach & Michael Traugott, eds. (Sage, 2007 forthcoming).

⁸ Heath, Fisher and Smith, “The Globalization of Opinion Research,” 2005.

⁹ Mattes, “Public Opinion in Emerging Democracies,” 2007 forthcoming.

However, it is less clear whether this logic applies to other institutions covered by the APRM questionnaire such as the Reserve Bank, or other areas such as corporate governance. It might be important to measure whether ordinary citizens see the private sector, especially big businesses as corrupt, and / or more or less corrupt than state agencies and elected officials. Beyond that, however, it is not clear what more citizens can really tell us about corporate governance.

There are also issues about which is it important to distinguish those who have had some experience with an institution or have heard of some issue. We can then examine in detail those who have had specific experiences because it tells us about the performance of the institutions (e.g. an experience of victimization by bureaucrats or elected officials). On the other hand, the lack of knowledge or experience may also be important to measure because it tells us how many citizens are being included in or excluded from key policy debates or access to public institutions.

But there are also a range of issues that fall under the purview of the APRM questionnaire on which the vast majority of citizens simply have too little experience to which to devote the scarce resources of a survey question item. The area of corporate governance, for example, springs to mind. Does it make sense to ask people who are shareholders in large corporation (themselves in all probability a very small minority) about things like their experiences in Annual General Meetings, or their knowledge of the company finances?

The APRM self review questionnaire should certainly provide the basis and overall guide for designing the questionnaire instrument for any public opinion survey questionnaire intended to support the APRM process. But it is not necessary to have ordinary people attempt to provide answers to the exact questions as posed in the APRM questionnaire, or even to each larger objective of the questionnaire. Quite simply, citizens cannot tell us about everything. Thus, we should not overload the survey questionnaire in an attempt to match the APRM instrument exactly. Rather, we need to decide what are the things that people can tell us (in terms of their experiences, awareness, behaviours, values, evaluations, or preference) that can help the review process complete the self review questionnaire?

Need for Clear Definition of Other Target Groups

In this respect, surveys of other representative samples of other populations, such as firms, civil society leaders, bureaucrats or technical experts may be appropriate tools to add into the tool boxes of the national self review. However, it is not clear that the APRM process has developed sufficiently clear definitions of each of these population groups. What are the defining characteristics of a firm, a civil society group, or a government official, let alone more ambiguous terms like “roleplayers” or “experts” that help us know who qualifies and who does not qualify. Only once we have such a working definition can we evaluate the representativeness of any attempt to sample these groups and begin to compare the views and experiences of each group with each other, with the citizens, and with their counterparts in other APRM countries.

Planning, Timing and Costs

Besides all the components we have discussed so far, conducting credible and quality citizen surveys takes advance planning in order to avoid rushing the complex issues of designing the survey instrument, as well as the sampling strategy. Last minute planning is likely to result in simply adopting existing questionnaires that might not be maximally appropriate for the purposes, and allowing survey companies or national statistic offices to impose their own standard operating procedures, which may also be inappropriate. While some of the steps presented below can be done in parallel, rather than sequentially, based on my experience in the Afrobarometer, country teams should anticipate at least a five to six month time span from the point of deciding to pursue the research to receiving usable results.

Questionnaire Design	4 weeks
Advertising and Awarding Bids To Research Provider	3 weeks
Questionnaire Translation	1 week

In House Pilot of Questionnaire and Redesign	2 weeks
Sample Design, Sample Drawing	2 weeks
Training Fieldworkers	2 weeks
Field Pilot	1 week
Fieldwork	4 weeks
Data Entry, Cleaning, Presentation of Marginal Results	4 weeks

Based on my experience, nationally representative surveys in Africa are expensive compared to other continents, and costs may vary widely depending on the size and infrastructure of the country, and whether one selects a for-profit or not-for-profit research firm. In general, national teams should anticipate spending anywhere between US\$85,000 to US\$125,000 for a survey of 1,200 respondents, again depending on the country and the selected fieldwork provider.

On the other hand, depending on the country in question, as well as when a given country wants to conduct their exercise, a significant amount of public opinion data covering a wide range of topics covered by the APRM (especially in the areas of socio-economic and political governance) already exists. The *Afrobarometer* has just finished its most recent round of surveys of nationally representative samples of citizens in 18 African countries, the largest survey project ever conducted on the continent, conducted in a twelve month window between March 2005 and February 2006. The countries included in this round were:

- West Africa, Benin, Cabo Verde, Ghana, Mali, Nigeria, Senegal
- East Africa Kenya, Madagascar, Tanzania, Uganda
- Southern Africa Botswana, Lesotho, Malawi, Mozambique, Namibia, South Africa, Zambia, Zimbabwe

In addition, we have now conducted 3 separate surveys each in 12 countries that provide the first evidence ever collected about trends spanning a six year period (circa 2000, circa 2003, circa 2005 in Botswana, Ghana, Lesotho, Malawi, Mali, Nigeria, Namibia, South Africa, Tanzania, Uganda, Zambia, Zimbabwe).

Finally, Afrobarometer plans on conducting new surveys in these 18 countries (with a few other possible additions) beginning in 2008. Single countries, or a group of countries, who plan on undergoing self review in the 2007 or 2008 may be able to obtain survey data at a far cheaper cost than if they initiated their own survey by contributing a portion of the Afrobarometer fieldwork costs, and/or paying for additional questions. More information about the project and contact details can be located at www.afrobarometer.org.